CHARACTERIZATION OF LIGHT-SIGNALING IN THE CIRCADIAN PHOTORECEPTOR PROTEINS VIVID AND CRYPTOCHROME

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Anand Teertha Vaidya

August 2012

CHARACTERIZATION OF LIGHT-SIGNALING IN THE CIRCADIAN PHOTORECEPTOR PROTEINS VIVID AND CRYPTOCHROME

Anand Teertha Vaidya, Ph. D.

Cornell University 2012

Biological clocks are ubiquitous in nature; observed in almost all life forms. The role of clock is not only to synchronize the various biological activities of an organism, but also to synchronize the organism to the day/night cycle and gain the most out of sunlight. At the molecular level, a negative feedback loop is the basis of eukaryotic clocks and light triggers activity within the feedback loop affecting either the negative or the positive arm of the loop. In the current study, both these modes of light input to the clock have been investigated.

Vivid (VVD) is a photoreceptor protein in the fungus *Neurospora crassa* and is important for photoadaptation in continuous light. VVD senses light via a flavin cofactor incorporated in a light, oxygen and voltage sensing (LOV) domain. Light induces the formation of a covalent bond between the flavin and a conserved cysteine, which causes conformational changes with VVD that lead to its dimerization. This dimer is the signaling state of VVD and it was crystallized in this study using a Met135I:Met165I variant that increases the lifetime of the light state. The structure reveals the mechanism of light-signaling where in light induced bond formation alters hydrogen bond pattern, leading to the release of the N-terminal loop of the protein, which ultimately docks into the other subunit to form the dimer. The dimeric structure was confirmed in solution and in the organism. Based on similarities in sequence alignments, this mechanism appears to be conserved among all fungal LOV proteins.

Cryptochrome (CRY) is the principal photoreceptor in the fruit fly *Drosophila melanogaster* that senses light via a flavin cofactor. Determination of the conformational changes associated with light exposure and the role of the flavin redox state in these conformational changes was the aim of the investigation. A trypsin based limited-proteolysis assay was developed to monitor the light state conformation of CRY. This assay in combination with mass spectral analysis, kinetic studies and mutational analysis reveals structural rearrangements in and around the C-terminal extension, which are strongly coupled to the light-induced redox changes of the flavin. Chemically reducing CRY in the dark was sufficient to replicate the structural rearrangements observed in the presence of light, which strongly indicates that the role of light is to reduce the flavin.

BIOGRAPHICAL SKETCH

Anand Teertha Vaidya was born in 1985 to Mrs. Parimala Vaidya and Dr. Vijaya Rao Vaidya in Hyderabad, a major city in the southern part of India. He is the youngest of three children and was raised with equal emphasis on science and religion/spirituality. He went to Saint Anthony's High School, Hyderabad up to the 10th standard, which he finished in 2000, and studied the two years of Intermediate (11th and 12th standard) at Little Flower Junior College, Hyderabad. He obtained his Bachelor's degree from Nizam College, Osmania University, Hyderabad, where he studied Mathematics, Physics, Chemistry, English and French. He enjoyed Prof. Ram Reddy's Organic chemistry, Prof. Uma Vuruputuri's Physical chemistry and Prof. Vijay Kumar's English poetry classes. During the three years of his Bachelor's, he did summer research with Prof. Susanta Mahapatra, University of Hyderabad, Hyderabad, Prof. Mangala Sunder Krishnan, Indian Institute of Technology Madras, Chennai and Prof. Srinivasan Chandrasekaran, Indian Institute of Science, Bengaluru, which deepened his interest for research in sciences. In 2005, he moved out of his hometown, to the northern part of India (at the foothill of the Himalayan Mountains), for his Masters in Chemistry at the Indian Institute of Technology, Roorkee. He specialized in Organic chemistry and worked with Prof. Ashok Kumar Singh. Having worked in theoretical chemistry and organic chemistry, he was drawn to biology as he was intrigued by human behavior and wanted to study his (human's) interactions with his surroundings. At Cornell, he worked with Prof. Brian Crane to understand the mechanisms of light sensing & signaling of photoreceptor proteins involved in the biological clocks.

To my parents

ACKNOWLEDGMENTS

Over the past five years at Cornell, there have been innumerable people who have contributed towards my Ph.D. – some directly and many others indirectly. My path has crossed those of various others and each of these intersections changed my course by various degrees, to bring me to where I am now. I would like to take this opportunity to thank all those who have contributed to my journey till now, some that I know and many that I don't.

I thank my advisor Brian Crane for his patience, freedom and guidance throughout the five years. I also learned a lot just by observing him work, be it his enthusiasm for science or his unending optimism or his ability to juggle various roles with aplomb. He is not only a great scientist, but also a fantastic person and is a perfect role model. I am fortunate to have worked with him. Joanne Widom and her work ethics have been a constant source of inspiration. Conversations with Alex Bilwes were always delightful. Former lab members Brian Z, Sudhamsu, Abiola, Mike, Bhumit and Jaya helped me get started and taught me all the basic stuff in the lab. I thank them for putting up with my 'neubness' and for helping me find my feet. Thanks to Brian Z for getting me started on the clock-project. Interactions with Sudhamsu and Abiola almost always made me think and those with Bhumit have been refreshingly sarcastic. I admire Mike's coolness, patience and his teaching abilities. Jaya, you have been a wonderful friend and I am grateful to you for helping me come out of my shell during the early days. Xiaoxiao has been a wonderful friend and a great lab mate. Her sense of humour, openness and insights into human interactions makes her an 'interesting' companion. Ria has been my 'go-to' person in the lab whenever I have some lab news or need a

wake-up call for lab meetings or just talk about the ‘heroes’, the ‘*ranis*’ and others. I thank Xiaoxiao and Ria for all the help and for putting up with my meanness. I was inspired by Tom’s lab note book and although it didn’t materialize, I wanted to maintain mine like that. I always enjoyed the arguments Tom and I had with Abiola, and was pumped by his comments about my beard. His opinion about the Crane lab was pivotal for my decision to join the lab and I thank him for all the interactions. Ken, his theatrical descriptions and hilarious one liners have been thoroughly entertaining. I admire his patience and motivation and thank him for all the philosophical discussions. Greg has been a fun company and a good friend. I acknowledge my fellow night-shifters, Ken and Greg, for all the trash-talk and for ‘consistently keeping it real’, because ‘success never sleeps.’ All the other member of the lab, Bee, Camille, Pete, Sarah, Karen, Anna, Gabby, Dipanjan, Craig, Dowoon, Poulami, Anupam, Ushati, Lindsey, Steven, Meagan, Erika and Nancy, have made the lab a wonderful place to work and I thank them for all the good times.

I thank my committee members Steve Ealick and Rick Cerione for all the discussions, inputs and help.

Amit and I have been house mates for the past four years and things have been surprisingly smooth between us. He helped me in various day-to-day practical aspects of life and I have taken a couple of important leaves out of his book on how to deal with people. I thank him for all the fun time together, for his kindness and for his willingness to help. I cherish my interactions with Shruti. Thanks are due to Aritro, who is ever ready to go out of the way to help. I enjoyed a lot of hearty conversations with him and will miss the masseur in him. Life here would have been mundane without them.

Kumar was my house mate during my first year at Cornell and has been my philosopher friend. I have gained a lot from his experiences and I am grateful for his company and his pieces of advice. I thoroughly enjoyed Shridhar's company, be it watching cricket, or watching Telugu films or meeting outside S.T.Olin and I acknowledge him for all the advice. I thank Anushree for the words of wisdom.

Raj has been a friend in the true sense and is a stimulating company. His philosophy about society, about movies and about the purity of language among many others, has had a profound impact on me. I am thankful to him for his concern, for the good company, for introducing me to Hindustani music and for my improved Hindi. Suresh is a superb person who keeps life simple and tries to be happy always. It has been great knowing him and I am thankful to him for making me a part of SPICMACAY. Vidhya and I became close friends in a short period of time, and I am grateful to her for making me rethink and realise about happiness and contentment, about my priorities and about innumerable other things. Vidhya, your company was delightful to say the least and thank you for the wonderful friendship.

Vivek and Pavithra have been my closest friends during the ups and downs, and also during the normal times. The level of comfort I share with you brings me immense joy, confidence and peace. Pavithra, your child-like enthusiasm, honesty and respect for others, makes you a gem of a person and I have realised a lot about the ins and outs of life from our conversations. Vivek, we talk about everything under the sun and beyond, and I almost always come out with a new perspective, be it science, philosophy or anything else. Your patience, tolerance and purity never cease to amaze me. The world needs more people like you both.

I shared food – both cooking and eating, with Raj, Suresh, Vivek, Amit, Siddharth, Shruti and Kumar and I would like to express my sincere gratitude to them for the good food, for making me a better cook, for the various discussions about everything and for bearing with my ‘dinner dialogues’ and my idiosyncrasies…the so called ‘short-circuit rules’.

I made some wonderful friends who have contributed to my well-being here and I would like to thank Lekha, Hari, Deepti, Ankush, Vidhya R, Parag, Debashree, Rachna, Siddarth, Poulami, Anupam, Ushati, Chinmayee, Stueti, Siddharth, Kasturi, Samanvaya, Tanay, Krishna, Sasi, Pranav, Manish and Srivatsav for the warmth and help. I would like to acknowledge the role of Cornell and the ‘gorges’ Ithaca for the serene environment to grow.

I am grateful to all my former teachers in India for their constant encouragement and support.

I am indebted to my family – my parents (Parimala and Vijaya Rao), my sisters (Srivathsa and Sravanti) and my grandmother (Shakuntala Bai), who form the cornerstone of my life, for all the sacrifices they made for me, for the values they instilled in me, and for the undiminishing love. I thank my uncle Jayathirtha Rao and brother-in-law Kiran Kulkarni for guidance and my relatives for their support.

I thank God for everything…for who I am, for what I have & for what all I can do.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

al-1 – albino-1

al-2 – albino-2

AMPK - Adenosine monophosphate-activated protein kinase

ARNT - Aryl hydrocarbon receptor nuclear translocator

As phot1 - *Avena sativa* phototropin1

ASQ - anionic semiquinone

At CRY - *Arabidopsis thaliana* cryptochrome

ATP - adenosine triphosphate

BMAL1 - Brain and muscle ARNT-like 1

BLUF - sensory proteins of blue-light using flavin

ccgs - clock controlled genes

CCM - C-terminal tail coupled motif

CCT - cryptochrome C-terminal

CK-1a - Casein kinase-1a

CK-II - Casein kinase II

CLK – CLOCK

CNS – Crystallography and NMR System

con-6 - conidiation-specific protein gene-6

con-10 - conidiation-specific protein gene-10

CPD - cyclobutane pyrimidine dimer

CRY – Cryptochrome

CRY-DASH - cryptochrome-*Drosophila*, *Arabidopsis*, *Synechocystis*, *Homo*

CTT - C-terminal tail

Cu-Phe - Cu-Phenanthroline

CWO - Clockwork Orange

CYC – CYCLE

DBT – Double time

DLZ - 6,7-dimethyl-8-(1'-D-ribityl) lumazine

*Dm*CRY – *Drosophila melanogaster* cryptochrome

DNA – Deoxyribonucleic acid

DSM - dark-state monomer

DTT – dithiothreitol

EDTA – ethylenediaminetetraacetic acid

EPR - electron paramagnetic resonance

FAD - Flavin adenosine diphosphate

FFC - FRQ:FRH complex

FMN - Flavin monophosphate

FRET - Förster resonance energy transfer

FRH – FRQ-interacting RNA helicase

FRQ - FREQUENCY

FWD-1 - F-box/WD-40 repeat-containing protein-1

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

*Hs*CRY – *Homo sapiens* cryptochrome

ISM – intermediate-state monomer

JET – JETLAG

LOV - light-oxygen-voltage sensing

LSD - light-state dimer

mCRY – mammalian cryptochrome

MPD – 2-methyl-2,4-pentanediol

MTHF - 5,10-methylenetetrahydrofolate

N-cap - N-terminal cap

NMR - nuclear magnetic resonance

PAR - proline and acidic amino acid–rich

PAS - Period-Aryl hydrocarbon receptor nuclear translocator-Single minded

PCR – polymerase chain reaction

PDP 1 ε/δ - PAR Domain Protein 1ε/δ

PEG – polyethylene glycol

PER – PEROID

PHD - photolyase homology domain

PKA - Protein kinase A

PP2A - Protein phosphatase 2A

PTFL - post-translational feedback loop

PYP - Photoactive yellow proteins

RNA – Ribose nucleic acid

*Rs*CRY - *Rhodobacter sphaeroides* cryptochrome

SAXS - small angle X-ray scattering

SGG – Shaggy

TIM – TIMELESS

TTFL - translational feedback loop

UV- ultraviolet

VRI – Vrille

VVD - VIVID

WC-1 - WHITE COLLAR-1

WC-2 - WHITE COLLAR-2

WCC – White Collar Complex (WC-1:WC-2 complex)

WT – wild type

Chapter 1

MOLECULAR BASIS OF *NEUROSPORA* AND *DROSOPHILA* CIRCADIAN CLOCKS

1.1 Introduction

Agni (meaning fire in Sanskrit) or energy is the basis of life, and therefore, existence. Sunlight is the most important source of energy for plants to make glucose (i.e. food) via photosynthesis and all animals depend on plants, directly or indirectly, for survival. The Earth's rotation and revolution confers a variation in the presence and absence of sunlight within a 24 hour cycle and also variations in the intensity of the light that are correlated to our seasons. Both of these affect the temperature on earth. Plants depend solely on the sun for energy and can sense day to night transitions, the length of daytime over various seasons, and the intensity and direction of light, to maximize photosynthesis. On the other hand, animals depend on the sun for many aspects of their lives, including timing their sleep/wake cycle, determining their mating season, visual inputs and synthesizing vitamins.

All the organisms that sense the sunlight have adapted to the variation in light intensity by developing a timekeeping mechanism or a biological clock. This clock puts the various biological activities of an organism in a ~24 hour rhythm. This 24 hour rhythm has evolved to coincide with the 24 hour time period of the earth's rotation, but does not depend on the sunlight for sustenance; it persists even in the dark and is called a 'free-running' clock. This free-running biological clock with a 24 hour time period is called a circadian clock (circadian means 'around a day' in Latin). The

circadian clock leads to anticipation of dawn and dusk and gears up the organism to get maximum benefit from the light. To gain maximum benefit, the clock must precisely predict the onset of dawn and dusk; therefore, the internal circadian clock must be set correctly to the day/night cycle. Ironically, resetting the clock depends on the periods of sunlight during the day/night cycle. Thus, sensing the sunlight resets an organism's circadian clock, thereby helping it to anticipate the sun precisely in the future, and gain the most out of sunlight.

For something to be called a 'clock', it should be rhythmic with a certain time period, should maintain the rhythm for a long time, and should be able to reset to the 'actual' time. The circadian clock satisfies/achieves these criteria and the rhythm in organisms arises from a molecular clock involving rhythmic synthesis and degradation of certain proteins. Maintaining equilibrium between synthesis and degradation requires a feedback loop, particularly a negative feedback loop. At the molecular level, the negative feedback loop is composed of interacting macromolecules (proteins, DNA and RNA) within the cell.

1.2 Molecular Basis of Biological Rhythms

In eukaryotes, the central feedback loop of the circadian clock is initiated by a protein (or a group of proteins) called the positive element, which activates transcription of certain genes (Figure 1.1). These genes can be divided into two types - the clock genes, which are involved in the feedback loop, and the clock controlled genes (*ccgs*), which control the rhythmic activities of the entire organism. The positive element(s) initiates transcription of the clock genes in the nucleus, resulting in the production of

mRNAs, which are transported to the cytoplasm and translated into proteins. These proteins, called the negative element(s), block the transcription activity of the positive element(s) and prevent their own synthesis, thus closing the loop. This inactivation of the positive element also stops the synthesis of the clock controlled proteins arising from the *ccgs*. Reducing levels of the negative element(s) starts a new cycle by resulting in the activation of the positive element(s). Thus, the cyclic increase and decrease in the levels of the negative control protein forms the basis of circadian rhythms of an organism. This is the transcriptional translational feedback loop (TTFL) model of circadian rhythms. Though the numbers of players and the complexity varies in different organisms, the TTFL model is the basis of the clock in plants, fungi and animals [1-6].

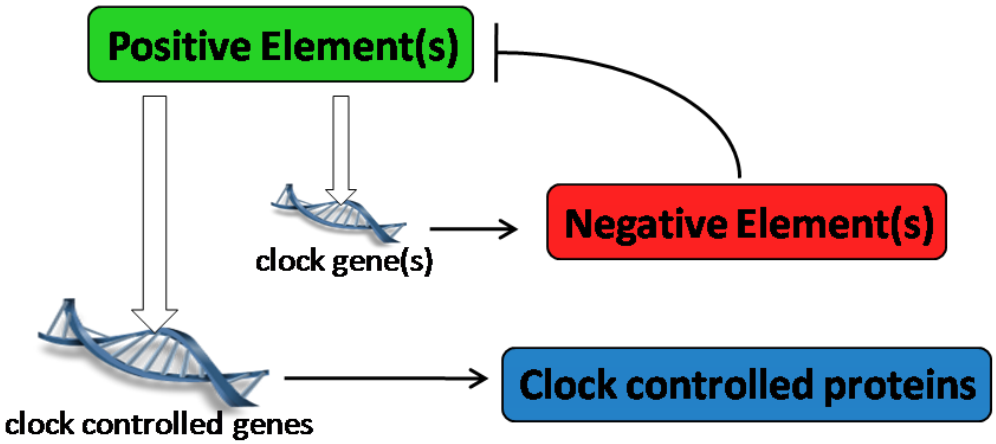


Figure 1.1 Negative feedback loop is the basis of a circadian clock. Negative element inhibits the activity of positive element, which inhibits the synthesis of negative element and clock controlled proteins. The clock controlled proteins determine the rhythm of various processes in the organism

Cyanobacteria are the only prokaryote with a known clock. This may be linked to the fact that they perform photosynthesis. Their clock is formed by rhythmic phosphorylation and dephosphorylation of the protein kaiC, which is controlled by the proteins kaiA and kaiB respectively. This is a post-translational feedback loop (PTFL) and is sufficient to drive the cyanobacterial clock even when they are growing in a test tube. Recently, a similar PTFL was discovered /observed in mammalian red blood cells, which lack the ability to perform transcription. It is important to note that the TTFL model relies extensively on post-translational modifications such as phosphorylation, acetylation, methylation, sumoylation and ubiquitination, some of which are rhythmic. It is not clear if the PTFL is another model for the clock, or if it a consequence of the clock, or if it is part of an oscillator yet to be discovered [7, 8].

Synchronizing the core clock with the day/night cycle involves sensing the sunlight. Light can influence the feedback loop by affecting either the positive element or the negative element. Light starts a new round of gene expression, either by activating the positive element or by deactivating the negative element. Blue-light, which has the highest energy and maximum penetration in water among the visible wavelengths, activates photosensory proteins that provide light input to the feedback loop.

The fruit fly *Drosophila melanogaster* and the fungus *Neurospora crassa* are model organisms to study the clock at the molecular level. The above described feedback mechanism was first proposed for the *Drosophila* clock and the negative feedback loop was first confirmed using the *Neurospora* system [7]. In *Neurospora*, the positive element senses blue-light and activates gene expression, while in *Drosophila*

the photosensor degrades the negative element that leads to activation of the positive element. Here, we study both types of photoreceptors from these organisms.

1.3 Components of the *Neurospora crassa* Circadian Clock

1.3.1 Introduction

Neurospora crassa is a type of red bread mold fungus and is used as a model organism to study the circadian clock. It is mainly found as a mycelium, a vegetative state containing connected fungi with cytoplasmic streaming, and under certain physical stimuli like light exposure and mechanical injury, they form asexual spores. These spores are called macroconidia (or conidia) and conidiation, which peaks just before dawn, is a rhythm-marker for the circadian clock [9]. Conidiation is a simple marker for the rhythm generated by the internal clock and has been instrumental in unraveling the molecular basis of the clock. The conidia are bright orange in colour due to carotenoids. The biosynthesis of carotenoids is under the control of the circadian clock and the intensity of the orange colour is used as a marker for photoadaptation.

1.3.2 Hands of the Neurospora clock

The *Neurospora crassa* circadian clock is a negative feedback loop at the molecular level (Figure 1.2). This loop involves the proteins WHITE COLLAR-1 (WC-1), WHITE COLLAR-2 (WC-2), FREQUENCY (FRQ) and FREQUENCY-INTERACTING RNA HELICASE (FRH) [10]. WC-1 and WC-2 form a WC complex (WCC) and bind to the promoter of the *frq* gene around late night/early day. This leads to the rise in the FRQ levels around mid to late day and results in FRQ binding to FRH. The FRQ:FRH complex (FFC) enters the nucleus and interacts with the WCC, which induces

phosphorylation of the WCC and FRQ by casein kinases. Hyperphosphorylated WCC is inactive, leading to its dissociation from the *frq* promoter and the termination of *frq* transcription around late evening. Hyperphosphorylated FRQ is ubiquitinated by F-box/WD-40 repeat-containing protein-1 (FWD-1) and is tagged for proteasomal degradation around late night, leading to a new round of transcription by the WCC. Thus the FFC inhibits its own synthesis and forms a negative feedback loop with a time period of ~22.5 hours. Along with *frq*, the WCC activates other genes called the *clock controlled genes* (*ccgs*) that lead to proteins that control the rhythmic behaviour of the fungus [4, 5].

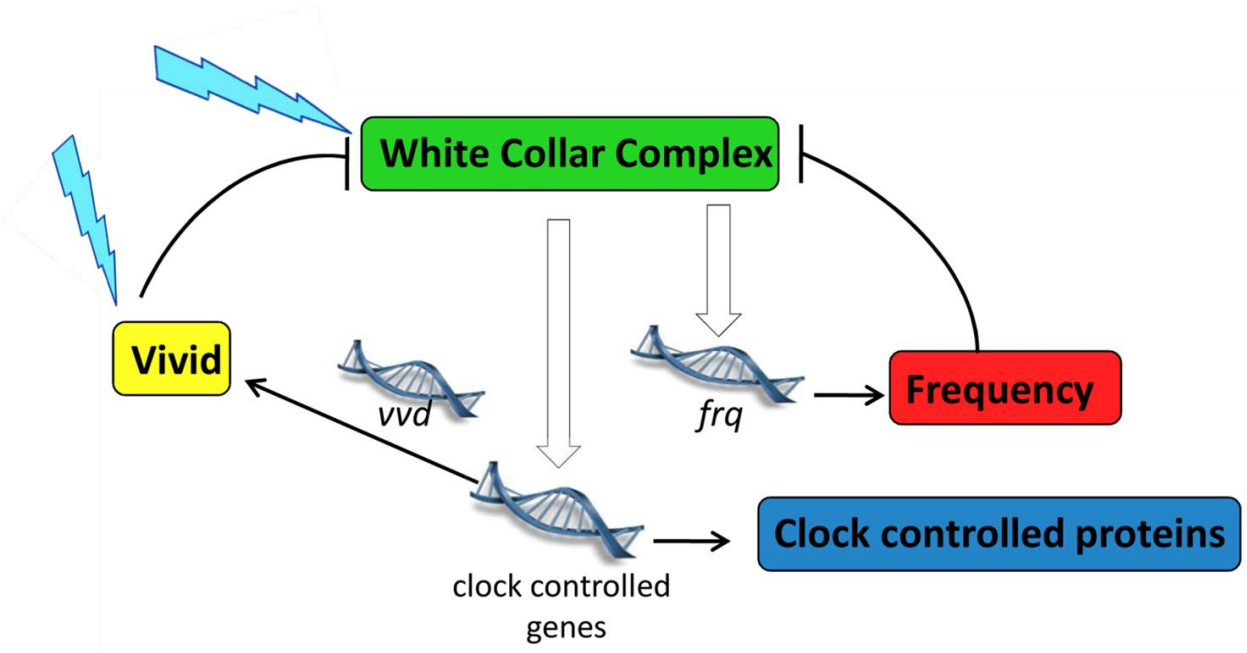


Figure 1.2 Components of the *Nerurospora* circadian clock. White Collar complex (WCC) is the positive element and FREQUENCY is the negative element. WCC is the primary photoreceptor that helps in photentrainment and VIVID, another photoreceptor, helps in photoadaptation.

1.3.3 Roles of phosphorylation and interlocking feedback loops

Phosphorylation state of clock proteins controls the stability, activity and cellular localization. FRQ is extensively phosphorylated at more than 75 residues and involves five kinases. Phosphorylation has opposing effects on the stability of the protein - most sites lead to degradation of FRQ, but the C-terminal sites stabilize it [11]. Casein kinase-1a (CK-1a) and casein kinase II (CK II) mediate most of the phosphorylation events; according to one model, CK-1a is the key kinase and can be considered a part of the FFC, along with FRQ:FRH [12]. Hypophosphorylated FRQ appears to shuttle between the nucleus and the cytoplasm while the hyperphosphorylated FRQ is predominantly localized in the cytoplasm [11]. Recently, it was shown that CK1a phosphorylates the N-terminus of FRQ *in vivo*, leading to a conformational change in the protein. This exposes a protected site for phosphorylation, which results in its degradation [13]. FRQ phosphorylation by CK II is implicated in the temperature compensation of *Neurospora*. FRQ is always associated with FRH and it is likely that FRH regulates phosphorylation of FRQ. The WC proteins in the WCC are phosphorylated by protein kinase A (PKA), CK-1a and CK II, facilitated by FRQ. Hypophosphorylated WCC is active and is localized in the nucleus. The inactive hyperphosphorylated WCC is predominantly found in the cytoplasm and phosphatase 4 (PP4) and phosphatase 2A reactivates it by dephosphorylation [12].

Neurospora crassa has other feedback loops that interlock with the core feedback loop, which could be important for the robustness and stability of the clock. FRQ activates *wc*-2 transcription and positively regulates WC-1 at a post-translational level. WC-1 and WC-2 form another feedback loop where WC-2 positively regulates

WC-1, while WC-1 negatively regulates *wc-2* via a repressor and activates *wc-2* in the presence of FRQ [5, 11]. The clock genes are also regulated at the post-transcriptional level [11].

1.3.4 Mechanism of photo-entrainment and photo-adaptation

This negative feedback loop is endogenous and it synchronizes with the day/night cycle via a photosensor. In *Neurospora*, WC-1 is the principal photoreceptor of the circadian clock, acting as a blue-light photoreceptor (Figure 1.2) [14-18]. Light exposure of WC-1 leads to dimerization/multimerization of the WCC and rapid binding to the FRQ promoter initiating transcription [16, 19, 20]. WCC rhythmically binds to *frq* in the dark, but light exposure at any time initiates WCC-induced FRQ transcription, essentially resetting the clock [20]. WC-1 is a multi-domain protein containing two transcriptional activation domains – one at each termini, a zinc finger DNA binding domain, a nuclear localization signal, two Period-Aryl hydrocarbon receptor nuclear translocator-Single minded (PAS) domains and a light-oxygen-voltage sensing (LOV) domain [18]. The PAS C domain of WC-1 has been implicated in dimerization with the PAS domain of WC-2, while the LOV domain (sometimes referred to as PAS A) is the light sensor [17, 21, 22]. Light-induced gene activation by the WCC can be divided into two groups based on the response time of activation – immediate light-inducible genes and late-responsive genes. The immediate light-inducible genes include *frq*, *wc*-1, vvd, *albino-1 (al-1*), *al-2*, *conidiation-specific protein gene-6* (*con-6)* and *con-10* [23]. The *al-1, 2* genes lead to the biosynthesis of carotenoids, the *con-6, 10* are important for conidiation and *vvd* is responsible for photoadaptation. The burst of gene expression after light exposure is down regulated by another LOV-photosensor, VIVID (VVD) and

this process is called photoadaptation (Figure 1.2) [24, 25]. Immediately after its synthesis, VVD enters the nucleus and down regulates WCC by interacting with the LOV domain of WC-1 [19, 26, 27]. VVD helps *Neurospora* respond to different intensities of light. Thus, VVD forms a negative feedback loop with the WCC and provides the clock with a molecular gate to control its sensitivity to light. VVD has a single LOV domain with a short N-terminal extension. The LOV domain of VVD and WC-1 are similar, and *Neurospora* strain in which the WC-1 LOV is switched with the VVD LOV is still functional [28]. VVD is also implicated in the temperature entrainment and compensation of *Neurospora* [29].

1.3.5 Structure and function of PAS domain

The PAS domain is a sensory motif in many signaling proteins and plays an important role in all kingdoms of life. PAS domains regulate phototropism in plants, circadian entrainment in animals, ion-channels in humans and nitrogen fixation in bacteria. Many of these diverse functions involve PAS-PAS interactions. A PAS domain is a simple mix of α/β type protein structure, with α-helices on either side of a central five-stranded anti-parallel β-sheet. One of the faces of the β-sheet forms a hydrophobic pocket with the α helical element and the other face, due to its hydrophobic nature, is often involved in protein:protein interactions. The domain topology involves two β-strands (Aβ, Bβ), followed by a series of short α-helices (Cα, Dα, Eα, Fα) and three anti-parallel B-strands (Gβ, Hβ, Iβ). Many PAS domains contain N-terminal or C-terminal extensions, which directly interact with the hydrophobic region of the core β-sheet. Some PAS domains bind co-factors such as flavin adenosine diphosphate (FAD), flavin monophosphate (FMN), heme, citrate, etc.; those binding flavins are classified as light-

oxygen-voltage sensing (LOV) domains. In response to an external stimulus such as chemical ligand, light or redox potential, the PAS domain undergoes conformational changes in the variable regions which ultimately allow interaction with new partners or effector domains [6, 30-32].

1.3.6 Flavin: Nature's blue light sensor

Nature employs two organic small molecules to sense blue light - p-coumaric acid and flavin. Flavin contains an isoalloxazine tricyclic ring and is a part of the vitamin B2, riboflavin. Flavo-proteins have either a flavin mononucleotide (FMN) or a flavin adenine dinucleotide (FAD) and both the proteins studied here have a FAD (Figure 1.3). Flavins (specifically, the isoalloxazine ring) are versatile redox molecules and can undergo both 1-electron and 2-electron redox chemistry. It can exist in four different redox states – oxidized, anionic semiquinone, neutral semiquinone and reduced, with different absorption spectrum. The oxidized state, which is yellow in color, is generally the ground state and absorbs blue light of ~450 nm wavelength. This reduces it to one of the three reduced states depending on the electronic environment around the flavin.

1.3.7 LOV domain: Mechanism of light-sensing and light-signaling

LOV domain is a PAS domain with a flavin cofactor and is found in archaea, bacteria, prostists, plants and fungi [33, 34]. The light sensing ability of LOV domains is well studied, but the mechanism of voltage and oxygen sensing are not well understood. Blue-light induces the formation of a new covalent bond between the C(4a) carbon of the flavin and the sulphur of a conserved cysteine of the protein (Figure 1.4). The kinetics of the light reaction has been studied up to the picosecond timescale, but the

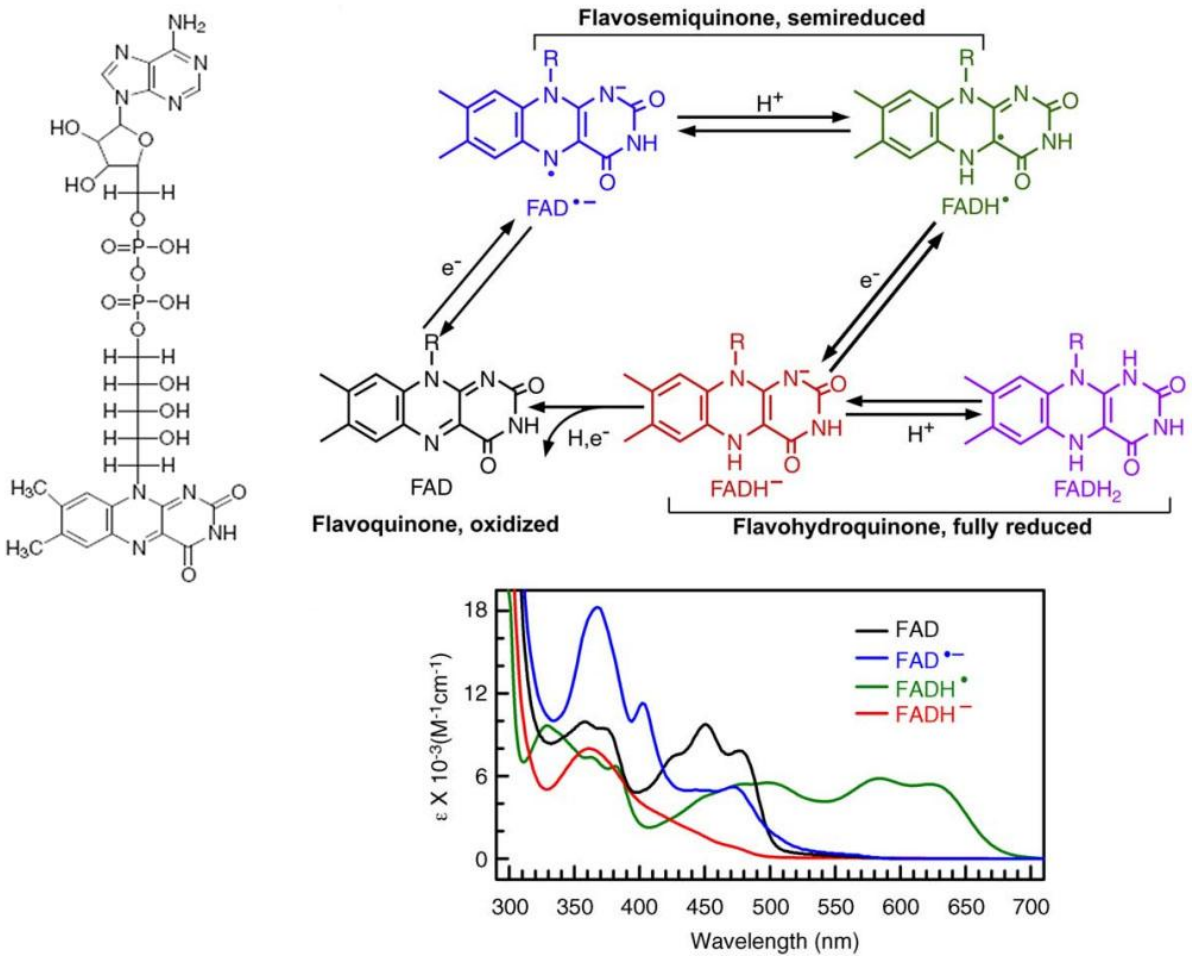


Figure 1.3 Redox states of FAD and their absorption spectra. FAD (Left) its redox states (right top) and the absorption spectrum of the redox states (bottom). The colour of the redox state corresponds to the colour of the spectrum. (Reproduced with permission of ELSEVIER LTD. in the format Journal via Copyright Clearance Center, #62488036, Current opinion in plant biology, 2010. **13**(5): p. 578-586)

dynamics of the adduct formation are yet to be deciphered [35]. Energy from blue-light excites the flavin to a singlet state within nanoseconds, followed by an intersystem crossing to a triplet state in picoseconds. This leads to adduct formation via two possible mechanisms – an ionic mechanism or a radical-pair mechanism. Quantum chemical calculations and low-temperature electron paramagnetic resonance (EPR)

studies support the radical pair mechanism, which involves a hydrogen transfer from the conserved cysteine to the N(5) nitrogen of the flavin, resulting in a $FADH^{o}$-CH2-S^{o} radical pair. The unpaired electron in the $FADH^{o}$ resides on its C(4a) atom and the $FADH^{o}$ converts to a singlet state followed by recombination of the radical-pair, forming an adduct between the conserved cysteine and the C(4a) of FAD. Protonation of the FAD and the electron transfer could be concerted [36]. The dark-state and the light-state can be easily distinguished based on their absorption spectrum. This adduct is cleaved in the dark and the lifetime of the adduct varies from a few seconds in plant LOV phototropins to a few hours in bacterial protein YtvA and fungal protein VVD [36, 37]. A base-catalyzed mechanism has been proposed for the adduct cleavage, but the identity of the base is still unknown [36, 37]. Mutating residues around the flavin alters the lifetime of the Cys-adduct by varying the electronic environment around the flavin and the solvent accessibility to the flavin [37].

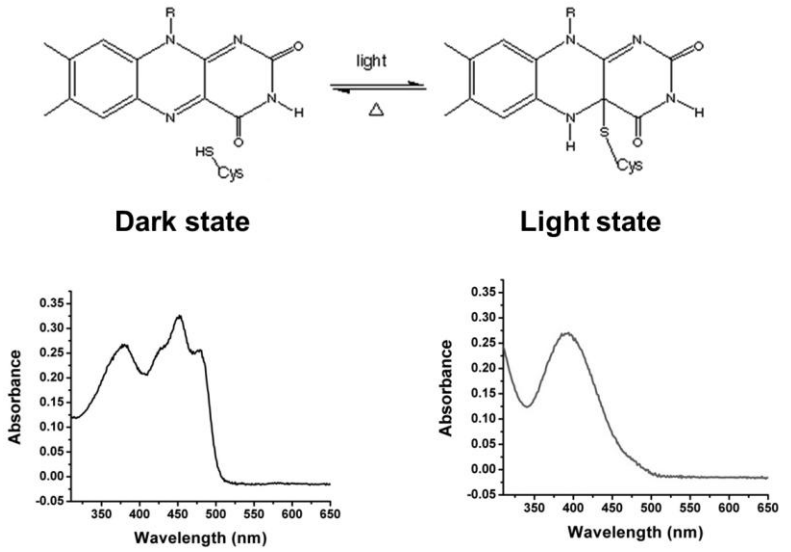


Figure 1.4 Photoreaction of LOV proteins. Light induced adduct formation in LOV domains and the absorption spectrum of the two states. Flavin is in the oxidized state in dark and in the reduced state in light.

Light-initiated conformational changes in LOV domains have been well studied in the plant LOV proteins, phototropins, the bacterial LOV protein YtvA and the fungal protein, VVD [31, 38, 39]. Phototropins have two N-terminal LOV domains followed by a conserved Jα helix and a Ser/Thr kinase domain. They control phototropism and regulate photosynthesis by light-induced autophosphorylation [40]. Structural analysis of just the LOV2 domain with the Jα helix reveals that the helix docks onto the LOV β-sheet and light releases the Jα helix by unwinding it [41]. The undocking of the Jα helix renders the kinase domain free to phosphorylate another phototropin molecule [40]. This light-induced undocking of the Jα helix is being used to control the spatio-temporal dynamics of various proteins in fusion with photoropin [42, 43], while other LOV proteins have been engineered to act as fluorescent proteins [44] and to place transcription under the control of blue-light [45]. VVD forms a homodimer in light, which is the signaling state of the protein; disrupting the dimer by mutating a key residue affects photoadaptation in *Neurospora* [31]. Exposing the dark-grown crystals of phototropin LOV2, YtvA and VVD to light reveals the light-induced molecular changes in the vicinity of the flavin [31, 38, 39]. Protonation of the N(5) in light alters the orientation of the side chain of a conserved glutamine residue by forming a new hydrogen bond. The newly flipped glutamine side chain forms another hydrogen bond with a residue on the other side of the glutamine. This shifts the N-terminal helix to a small extent in VVD, but structural constraints induced by the crystal prevent large conformational changes in both the systems [31, 46, 47]. Small angle X-ray scattering (SAXS) data on VVD hint at changes in the N-terminus due to light exposure [48]. The mechanism of light-induced

alterations in the global protein structure caused by the adduct formation is not known in phototropin LOV2 or VVD or any other LOV protein.

In chapter 2, we report the crystal structure of a fully light-adapted dimer of VVD. Based on *in vitro* and *in vivo* mutational studies coupled with *in silico* sequence analysis, we propose a mechanism for light-induced oligomerization, which is conserved in all fungal LOV proteins (including the *Neurospora* WC-1).

1.4 Components of the *Drosophila melanogaster* Circadian Clock

The *Drosophila melanogaster* circadian clock is a negative feedback loop at the molecular level involving many proteins (Figure 1.5) [4, 49]. CLOCK (CLK) and CYCLE (CYC) are transcription activator proteins that form a heterodimer and bind to an E-box (CACGTG) in the *Drosophila* genome at midday [50]. The E-box is in the promoter region of many clock-genes, including *period* (*per*) and *timeless* (*tim*), which are central to the clock [51, 52]. The proteins PER and TIM, along with the kinase Double-Time (DBT), enter the nucleus where the PER-DBT-TIM complex accumulates [53-55]. This interacts with the CLK:CYC complex, via PER and CLK, and leads to the phosphorylation of CLK. Hyperphosphorylation of CLK prevents the CLK:CYC complex from binding DNA and shuts down *per* and *tim* transcription [50, 56]. Thus, PER inhibits its own synthesis, completing the negative feedback loop and conferring rhythmicity to TIM and PER expression. PER cycling thus comprises the core feedback loop found in *Drosophila*.

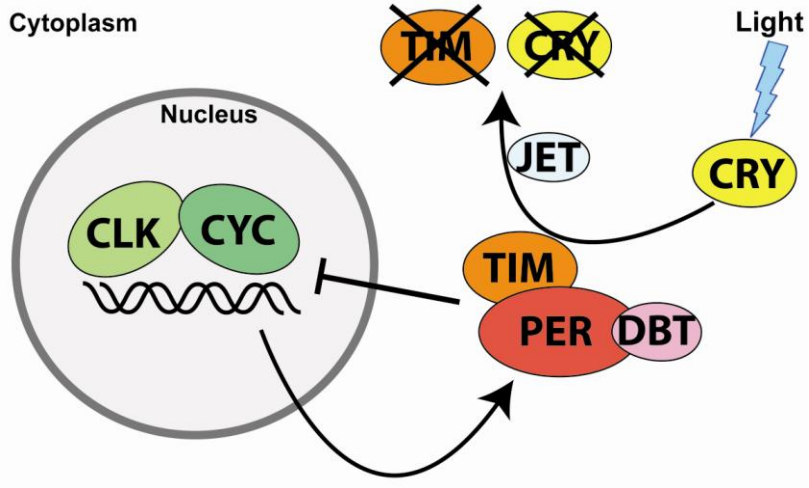


Figure 1.5 The negative feedback loop of the *Drosophila* circadian clock. CLK:CYC form the positive element, while TIM:PER:DBT form the negative element. CRY is the principal photoreceptor for photoentrainment.

1.4.1 Roles of phosphorylation and interlocking feedback loops

Phosphorylation plays a major role in the stability, activity and cellular localization of the clock-proteins. Phosphorylation of PER leads to its degradation, and this process is inhibited both by protein phosphatase 2A (PP2A) and by the PER-TIM interaction. PER phosphorylation by Casein kinase 2 (CK2) and TIM phosphorylation by Shaggy (SGG) tags them for nuclear entry [50, 51]. Phosphorylated TIM is recognized by the E3 ubiquitin ligase, Jetlag (JET), which ubiquitinates it for degradation [51, 57, 58].

In *Drosophila*, other feedback loops that interlock with the core loop have been identified. These loops involve the transcription factors Vrille (VRI), PAR Domain Protein 1ε/δ (PDP 1 ε/δ) and Clockwork Orange (CWO), whose corresponding genes are activated by the CLK:CYC complex [51]. VIR and PDP 1 ε/δ control the transcription of *clk* by inhibiting and promoting it respectively [54], while CWO inhibits the transcription activity of CLK:CYC by competing for the E-box [50, 51]. Though the roles of these

interlocking loops is not understood well, they seems to control the rhythmicity of output genes and could play a role in fine tuning the core loop.

1.4.2 Mechanism of light input via cryptochrome

Light input to the clock is mediated by the photoreceptor cryptochrome (CRY) (Figure 1.5) [50, 59-61]. The *cry* null mutants lose most of the light responses and are arrhythmic in constant light [62, 63]. CRY binds TIM in the presence of light and leads to phosphorylation of TIM. Phosphorylated TIM is recognized by JET and tagged for degradation [58, 64, 65]. As TIM levels drop, its binding interaction with PER is severed, leading to PER hyperphosphorylation and finally to PER degradation. PER degradation resets the clock and this light induced resetting is mediated by CRY. Fine tuning *Drosophila* light perception involves different isoforms of TIM. The short isoform is 23 residues shorter than the long isoform and is very sensitive to light because it binds strongly to CRY [50]. Recently, another protein, QUASIMODO, has been shown to provide CRY-independent light signal to the clock [66]. The molecular basis for temperature entrainment is not clearly understood, but PER could have a role in it. Alternate splicings of a *per* intron at different temperatures has been observed [50].

1.5 Photolyase: Blue Light Sensing DNA Repair Protein

1.5.1 Introduction

Cryptochromes have evolved from photolyases, which are DNA repair proteins that repair UV-damaged DNA. UV radiation damages DNA by crosslinking adjacent pyrimidine bases and resulting in two types of lesions – 1) the cyclobutane pyrimidine dimer (CPD) and 2) the 6,4-photoproduct. Though the repair mechanism for the CPD

and 6-4 lesions is similar, there are specific proteins for each – the CPD photolyase and the (6-4)PP photolyase. Though photolyases are found in all three kingdoms of life, they are absent in placental mammals (including humans) and the (6-4) photolyase is yet to be found in prokaryotes [62]. Photolyases have also been found in viral DNA. In mammals, damaged DNA is repaired exclusively via the excision repair pathways, by a light independent mechanism, using energy from adenosine triphosphate (ATP). Photolyase repair these lesions by using the energy from blue light (350 nm – 450 nm) via its 5,10-methylenetetrahydrofolate (MTHF) and flavin adenine dinucleotide (FAD) cofactors.

1.5.2 Mechanism and kinetics of DNA repair

The mechanism of DNA repair by photolyases has been well characterized both structurally and spectroscopically [61, 67-70]. Damaged DNA binds to photolyase, in dark, in a positively charged groove and the lesion is flipped out from the duplex, placing it close to the FAD. In the presence of light, FAD absorbs at 450 nm and is reduced to $FADH^{\cdot}$, with a conserved Trp-triad providing the electrons. The source of the proton is not clearly known. MTHF absorbs 350 nm light and transfers this energy to $FADH^{\cdot}$ to form the excited state $FADH^{\cdot *}$, which then repairs the lesion by a cyclic electron transfer. It takes a little more than 1 ns for the CPD repair and about 11 ns for the (6-4)PP repair [69, 70]. In CPD repair, the electron is transferred from the $FADH^{\cdot *}$ to the lesion in 250 ps followed by the repair in about 100 ps and the electron moves back to the $FADH^{o}$ in 700 ps. There is an unproductive back electron transfer from the lesion to $FADH^{o}$ before the repair, and its rate is currently under debate – one study estimates the electron is transferred in 350 ps while the other estimates 2.4 ns [71-74]. In the (6-

4)PP repair, the forward electron transfer from the $FADH^{-*}$ to the lesion takes 225 ps and the back electron transfer is faster, within 50 ps. A proton transfers from a conserved histidine in 425 ps, which leads to the repair of the DNA. This step is followed by the return of a proton and an electron in 10 ns. Once the proton is transferred to the 6-4PP, there is no back electron transfer and the repair goes to completion [69]. The competition between the fast back electron transfer and the slower proton transfer explains the low quantum yield of ~0.1 compared to ~0.9 of CPD repair [69].

1.6 Cryptochrome: Classification and Functions

1.6.1 Introduction

Cryptochromes are photolyase like proteins with no DNA repair properties. The term cryptochrome was used to refer to an unknown or 'cryptic' plant photoreceptor that imparts blue-light induced photoresponses. Two proteins were identified in plants and referred to as cryptochromes [62]. Subsequently, photolyase-like proteins were found in the human genome, but had no DNA repair properties and were classified as cryptochromes. Cryptochromes are less widespread than photolyases, but are found in all three kingdoms of life and seem to have evolved from two different precursors [61, 67]. The plant cryptochromes and their homologues have evolved from the CPD photolyase, while the animal cryptochromes and their homologues have evolved from the (6-4) photolyase. Plant cryptochromes are blue light photo-sensors, whereas the animal cryptochromes are divided into two groups – type I are photoreceptors and type II are not. All cryptochromes have a FAD, but the presence and identity of a second

cofactor is still under debate [75]. Cryptochromes show 25-40% sequence identity to photolyase, and have a photolyase homology domain (PHD) and a C-terminal extension of varying lengths (Figure 1.6) [62]. The cryptochrome C-terminal (CCT) mediates signaling and is implicated in the interactions with binding partners [61]. In addition to the plant and animal cryptochromes, there is a third class called the cryptochrome-*Drosophila*, *Arabidopsis*, *Synechocystis*, *Homo* (CRY-DASH), with unknown function.

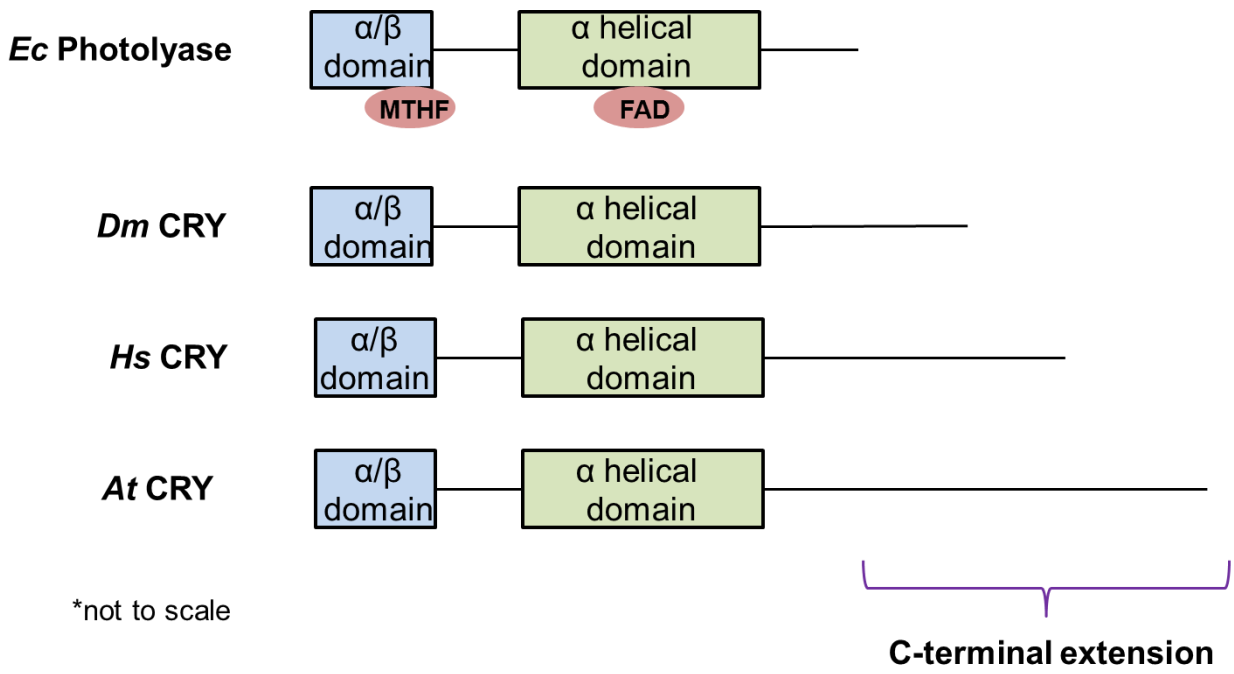


Figure 1.6 Domain architecture of Photolyase/cryptochrome family with varying C-terminal extension. *Ec* - *Escherichia coli*, *Dm* - *Drosophila melanogaster* (type I animal CRY), *Hs* - *Homo sapien* (type II animal CRY) and *At* - *Arabidopsis thaliana* (plant Cry).

CRY-DASH proteins bind to DNA, but unlike photolyase, they do not have modest ability to repair DNA and are instead implicated in signaling. They evolved separately from CPD photolyase and are proposed to be an intermediate form between photolyase

and cryptochrome [61]. Recently, the structure of a cryptochrome from *Rhodobacter sphaeroides* (*Rs*CryB) has been determined and it has three co-factors – FAD, a [4Fe-4S] cluster, and a 6,7-dimethyl-8-(1'-D-ribityl) lumazine (DLZ). The *Rs*CryB has homologs in proteobacteria and in cyanobacteria, which form a new subclass called the CryPro [76].

1.6.2 Plant cryptochromes

Blue-light responses in plants have been well studied in *Arabidopsis thaliana* (*At*), which has three cryptochromes – *At*Cry1, *At*Cry2 and *At*Cry3. *At*Cry3 is a DASH-type cryptochrome whose function is not well understood, while *At*Cry1 and *At*Cry2 have been the torch-bearers of the plant cryptochromes (and their homologues) and have been well characterized. These proteins are involved in controlling the plant's circadian clock by regulating flowering time, root development, stomatal opening, stress response, and photosynthesis [61, 77]. They have an N-terminal photolyase homology region involved in photoactivation and in homodimerization, and a C-terminal extension that is involved in signaling and binds to downstream targets [61]. Based on *in vitro* and *in vivo* studies, the FAD in *At*Cry1, 2 is in the oxidized state in the dark and light illumination reduces it to a neutral semiquinone ($FADH^{o}$) followed by re-oxidation in the dark over several minutes. Formation of the $FADH^{o}$ state results in conformational changes in the protein, specifically in the CCT, but a light dependent interaction with signaling partners via the CCT is not yet confirmed. Light sensitive phosphorylation has been observed both *in vitro* and *in vivo*, but a biological significance of phosphorylation, if any, is unknown [61, 78]. ATP binds to *At*Cry in the dark and in light it increases the lifetime of the $FADH^{o}$ radical [79]. Autophosphorylation of the N-terminal domain has

been observed *in vitro*, while the C-terminal domain is phosphorylated by an unknown kinase *in vivo* [61].

1.6.3 Type II animal cryptochromes

Type II animal cryptochromes are non-photoreceptor proteins involved in the core feedback loop of the circadian clock. The mammalian cryptochromes (mCRY1 and mCRY2) are the best investigated type II cryptochromes. Similar to the *Drosophila* clock, the mammalian circadian clock is a negative feedback loop with BMAL1 and CLOCK acting as the transcription activators that express *cry1*, *cry2*, *per1*, *per2* and *per3* genes. The CRY1, 2 and PER1, 2 proteins repress the BMAL1:CLK complex and complete the negative loop (Figure 1.7) [4, 80]. Recently, a crystal structure of BMAL1:CLK complex with the PAS domains and the DNA binding domains has been determined [43, 45]. The phosphorylation, acetylation and ubiquitination status of CRY determines its stability and spatio-temporal localization [61]. Mammalian CRY is expressed in all tissues [81] and could play an important role in peripheral clocks by serving as an energy sensor in liver tissues by activating adenosine monophosphate-activated protein kinase (AMPK) during glucose deprivation, which phosphorylates CRY1 and destabilizes it. This disrupts circadian rhythms in mouse liver cells [82]. Recently, it has been shown that CRY1, 2 are involved in regulating glucose metabolism by interacting with hormone receptors (specifically, glucocorticoid receptors) [83]. Though some studies have shown photoreceptor properties for the mammalian cryptochromes in living cells and for their homologs in insects, currently, their function *in vivo* is predominantly light-independent [84, 85]. Recent *Hs*CRY1 was purified *in vitro*

with the FAD mostly in an anionic semiquinone form, which did not reoxidize in dark [86].

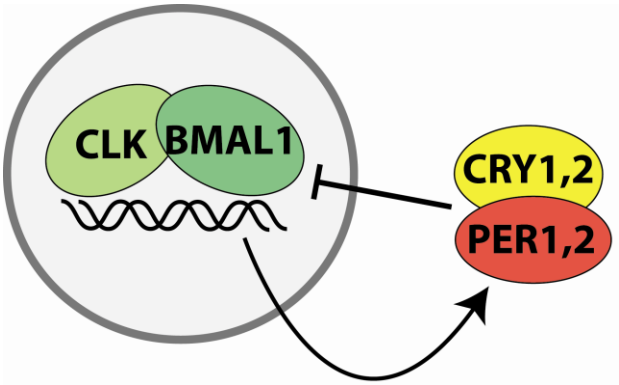


Figure 1.7 The core loop of the mammalian circadian clock. CLK:BMAL1 form the positive element, while CRY:PER form the negative element. CRY is not a photoreceptor in this system.

1.6.4 Type I animal cryptochrome: Drosophila cryptochrome

Drosophila cryptochrome is a well-studied type I animal cryptochrome. It is the principle photoreceptor in the *Drosophila* circadian clock and was identified by defects in behavioral rhythms in the cry^{b} mutants [59, 87]. *Dm*Cry has an N-terminal photolyase homology region that is important in photoactivation and in binding to TIM, and a C-terminal extension (CCT) that regulates photosensitivity [88]. TIM is phosphorylated during its association with CRY, which is then recognized by JET and tagged for degradation. As TIM levels go down, CRY is targeted by JET and leads to CRY degradation by ubiquitinating the CCT [58]. CRYΔ also lacking the CCT results in light-independent degradation of TIM and is the constitutively active form [89]. Thus, the CCT prevents TIM-CRY interaction in dark and light-excitation of the PHD promotes the interaction. A combination of *in vivo* EPR studies and *in vivo* spectroscopic analysis

show the oxidized state of FAD as the ground state and the anionic semiquinone (FAD^{o}) as the light excited state [84]. *In vitro* analysis showed that the electron transfer pathway for the light-induced reduction to be via three tryptophan residues (Trp-triad), conserved in the photolyase/cryptochrome family. But, disruption of the Trp-triad did not affect flavin reduction *in vivo*, hinting at an alternate electron transfer pathway [84, 90]. An alternate model of light sensing in the cryptochromes proposes a semiquinone ground state and an excited semiquinone as the light state [90-93]. Phosphorylation of *Dm*Cry has not been observed [94]. Recently, it was shown that cryptochrome regulates neuronal firing in a light-dependent manner [95]. In addition to its function as a photosensor, *Dm*Cry has been shown to mediate light-independent functions in peripheral clock in *Drosophila* [96] and also magnetosensitivity, which is dependent on <420 nm light [97-99].

1.6.4.1 Crystal structure of Drosophila cryptochrome

Recently, we determined the structure of *Dm*Cry with the C-terminal extension (Figure 1.8) [100]. The structure is similar to photolyase with N-terminal α/β domain and the C-terminal helical domain, which form the photolyase homology domain (PHD). The C-terminal 10 residues, called the C-terminal tail (CTT), are connected to the PHD via a 25 residue linker, which together form the C-terminal extension. The CTT occupies a groove analogous to DNA in photolyase and the Phe534 residue in the CTT is in the same position as DNA lesion in photolyase (Figure 1.9). Alanine substitution of the entire CTT, or just the FFW motif in the CTT, affects the stability of CRY and its interactions with TIM in cell cultures. The CTT is ~5 Å from the FAD and is thus optimally placed to respond to the redox changes of the FAD. The CTT and the linker

connecting it to the PHD are flanked by four loops which are specific to type I cryptochromes - the C-terminal lid (residues 420-446), the protrusion motif (288-306), the phosphate binding loop (249-263) and the CTT base loop (154-160). These motifs could be important for the light responses of the CTT.

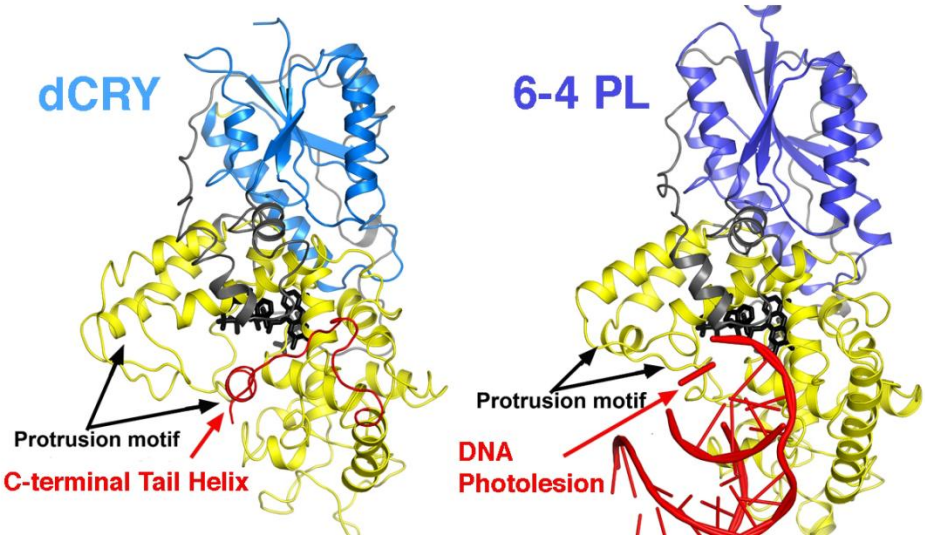


Figure 1.8 Structural similarities between cryptochrome and photolyase. *Dm*Cry (left) and *Dm*(6-4) photolyase (right) have similar 3-D structure and the C-terminal tail in CRY occupies the same groove as the DNA in photolyase (Nature, 2011, **480**(7377), 396-399).

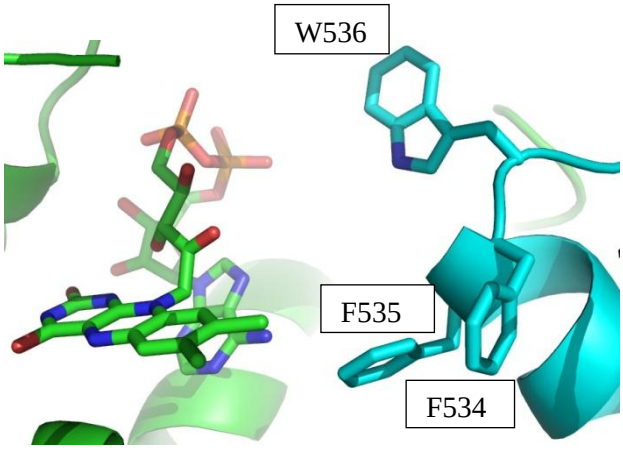


Figure 1.9 The FFW motif in the C-terminal tail of *Dm*Cry is in close proximity of the FAD.

The residue Cys416 in *Dm*Cry is within hydrogen bond distance to the N5 and O4 of the FAD, while the analogous residue in photolyase is an Asn residue which forms a hydrogen bond with protonated N5 of a reduced FAD. This explains why *Dm*Cry is light-excited to an unprotonated anionic semiquinone (FAD^{o-}) instead of the other protonated reduced states, $FADH^{o}$ and $FADH^{-}$. Phosphorylation of Thr518, in the C-terminal extension, had been observed in the crystal structure and in mass-spectral studies. The crystal structure lacked the MTHF antenna cofactor and isothermal titration calorimetry analysis showed no binding between *Dm*Cry and MTHF.

In chapter 3, we investigate the conformational changes associated with the CTT and the surrounding loops in response to light using mass-spectroscopy and mutational analysis. We develop an assay to follow CRY light-state conformation and investigate the correlation between the redox state of the FAD and the conformational changes. We propose a mechanism for light-signaling in *Drosophila* cryptochrome.

REFERENCES

1. Roenneberg, T. and M. Merrow, *Circadian systems: different levels of complexity.* Philosophical Transactions of the Royal Society-Ser B-Biological Sciences, 2001. **356**(1415): p. 1687-1696.
2. Pegoraro, M. and E. Tauber, *Animal clocks: a multitude of molecular mechanisms for circadian timekeeping.* Wiley Interdisciplinary Reviews: RNA, 2011. **2**(2): p. 312-320.
3. Jolma, I.W., et al., *Circadian oscillators in eukaryotes.* Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 2010. **2**(5): p. 533-549.
4. Young, M.W. and S.A. Kay, *Time zones: a comparative genetics of circadian clocks.* Nature Reviews Genetics, 2001. **2**(9): p. 702-715.
5. Heintzen, C. and Y. Liu, *The Neurospora crassa circadian clock.* Advances in Genetics, 2007. **58**: p. 25-66.
6. Vaidya, A.T., *The Chemistry Behind Circadian Clocks.* CHEMCOS: e-Journal of The Chemical Society, IIT Delhi, 2009(IV).
7. Brown, S.A., E. Kowalska, and R. Dallmann, *(Re) inventing the Circadian Feedback Loop.* Developmental cell, 2012. **22**(3): p. 477-487.
8. Mackey, S.R. and S.S. Golden, *Winding up the cyanobacterial circadian clock.* Trends in microbiology, 2007. **15**(9): p. 381-388.
9. Baker, C.L., J.J. Loros, and J.C. Dunlap, *The circadian clock of Neurospora crassa.* FEMS microbiology reviews, 2012.
10. Linden, H. and G. Macino, *White collar 2, a partner in blue-light signal transduction, controlling expression of light–regulated genes in Neurospora crassa.* The EMBO journal, 1997. **16**(1): p. 98-109.
11. Jinhu, G. and L. Yi, *Molecular mechanism of the Neurospora circadian oscillator.* Protein & Cell, 2010. **1**(4): p. 331-341.
12. Diernfellner, A.C.R., et al., *Phosphorylation modulates rapid nucleocytoplasmic shuttling and cytoplasmic accumulation of Neurospora clock protein FRQ on a circadian time scale.* Genes & development, 2009. **23**(18): p. 2192-2200.

13. Querfurth, C., et al., *Circadian Conformational Change of the Neurospora Clock Protein FREQUENCY Triggered by Clustered Hyperphosphorylation of a Basic Domain.* Molecular cell, 2011. **43**(5): p. 713.
14. Harding, R. and W. Shropshire Jr, *Photocontrol of carotenoid biosynthesis.* Annual Review of Plant Physiology, 1980. **31**(1): p. 217-238.
15. Crosthwaite, S.K., J.C. Dunlap, and J.J. Loros, *Neurospora wc-1 and wc-2: transcription, photoresponses, and the origins of circadian rhythmicity.* Science, 1997. **276**(5313): p. 763-769.
16. Froehlich, A.C., et al., *White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter.* Science's STKE, 2002. **297**(5582): p. 815.
17. He, Q., et al., *White collar-1, a DNA binding transcription factor and a light sensor.* Science's STKE, 2002. **297**(5582): p. 840.
18. Linden, H., *A white collar protein senses blue light.* Science, 2002. **297**(5582): p. 777-778.
19. Malzahn, E., et al., *Photoadaptation in Neurospora by competitive interaction of activating and inhibitory LOV domains.* Cell, 2010. **142**(5): p. 762-772.
20. Crosthwaite, S.K., J.J. Loros, and J.C. Dunlap, *Light-induced resetting of a circadian clock is mediated by a rapid increase in frequency transcript.* Cell, 1995. **81**(7): p. 1003-1012.
21. Cheng, P., et al., *WHITE COLLAR-1, a Multifunctional NeurosporaProtein Involved in the Circadian Feedback Loops, Light Sensing, and Transcription Repression of wc-2.* Journal of Biological Chemistry, 2003. **278**(6): p. 3801-3808.
22. Cheng, P., et al., *PAS domain-mediated WC-1/WC-2 interaction is essential for maintaining the steady-state level of WC-1 and the function of both proteins in circadian clock and light responses of Neurospora.* Molecular and cellular biology, 2002. **22**(2): p. 517-524.
23. Kozma-Bognár, L. and K. Káldi, *Synchronization of the fungal and the plant circadian clock by light.* ChemBioChem, 2008. **9**(16): p. 2565-2573.
24. Heintzen, C., J.J. Loros, and J.C. Dunlap, *The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting.* Cell, 2001. **104**(3): p. 453-464.

25. Shrode, L.B., et al., *vvd is required for light adaptation of conidiation-specific genes of Neurospora crassa, but not circadian conidiation.* Fungal Genetics and Biology, 2001. **32**(3): p. 169-181.
26. Hunt, S.M., et al., *VIVID interacts with the WHITE COLLAR complex and FREQUENCY-interacting RNA helicase to alter light and clock responses in Neurospora.* Proceedings of the National Academy of Sciences, 2010. **107**(38): p. 16709-16714.
27. Chen, C.H., et al., *Physical interaction between VIVID and white collar complex regulates photoadaptation in Neurospora.* Proceedings of the National Academy of Sciences, 2010. **107**(38): p. 16715-16720.
28. Cheng, P., et al., *Functional conservation of light, oxygen, or voltage domains in light sensing.* Proceedings of the National Academy of Sciences, 2003. **100**(10): p. 5938.
29. Hunt, S.M., et al., *The PAS/LOV protein VIVID controls temperature compensation of circadian clock phase and development in Neurospora crassa.* Genes & development, 2007. **21**(15): p. 1964.
30. Möglich, A., R.A. Ayers, and K. Moffat, *Structure and signaling mechanism of Per-ARNT-Sim domains.* Structure, 2009. **17**(10): p. 1282-1294.
31. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid.* Science's STKE, 2007. **316**(5827): p. 1054.
32. Losi, A. and W. Gärtner, *The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors.* Annual review of plant biology, 2011.
33. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Science's STKE, 2011. **4**(184): p. ra50.
34. Herrou, J. and S. Crosson, *Function, structure and mechanism of bacterial photosensory LOV proteins.* Nature Reviews Microbiology, 2011. **9**(10): p. 713-723.
35. Zoltowski, B.D. and K.H. Gardner, *Tripping the Light Fantastic: Blue-Light Photoreceptors as Examples of Environmentally Modulated Protein-Protein Interactions.* Biochemistry, 2010.

36. Kennis, J.T.M. and M. Alexandre, *Mechanisms of light activation in flavin-binding photoreceptors.* Flavins: Photochemistry and Photobiology, 2006: p. 287-319.
37. Zoltowski, B.D., B. Vaccaro, and B.R. Crane, *Mechanism-based tuning of a LOV domain photoreceptor.* Nature chemical biology, 2009. **5**(11): p. 827-834.
38. Marley, J., M. Lu, and C. Bracken, *A method for efficient isotopic labeling of recombinant proteins.* Journal of biomolecular NMR, 2001. **20**(1): p. 71-75.
39. Halavaty, A.S. and K. Moffat, *N-and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa.* Biochemistry, 2007. **46**(49): p. 14001-14009.
40. Christie, J.M., *Phototropin blue-light receptors.* Annu. Rev. Plant Biol., 2007. **58**: p. 21-45.
41. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch.* Science's STKE, 2003. **301**(5639): p. 1541.
42. Crane, B.R., L.M. Siegel, and E.D. Getzoff, *Structures of the siroheme-and Fe4S4-containing active center of sulfite reductase in different states of oxidation: heme activation via reduction-gated exogenous ligand exchange.* Biochemistry, 1997. **36**(40): p. 12101-12119.
43. Crane, B.R., *Nature's Intricate Clockwork.* Science Signalling, 2012. **337**(6091): p. 165.
44. Christie, J.M., et al., *Structural Tuning of the Fluorescent Protein iLOV for Improved Photostability.* Journal of Biological Chemistry, 2012. **287**(26): p. 22295-22304.
45. Huang, N., et al., *Crystal Structure of the Heterodimeric CLOCK: BMAL1 Transcriptional Activator Complex.* Science Signalling, 2012. **337**(6091): p. 189.
46. Crosson, S. and K. Moffat, *Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch.* The Plant Cell Online, 2002. **14**(5): p. 1067-1075.
47. Fedorov, R., et al., *Crystal structures and molecular mechanism of a light-induced signaling switch: the Phot-LOV1 domain from Chlamydomonas reinhardtii.* Biophysical journal, 2003. **84**(4): p. 2474-2482.

48. Lamb, J.S., et al., *Illuminating solution responses of a LOV domain protein with photocoupled small-angle X-ray scattering.* Journal of molecular biology, 2009. **393**(4): p. 909-919.
49. Rosbash, M., et al. *Transcriptional feedback and definition of the circadian pacemaker in Drosophila and animals*. 2007. Cold Spring Harbor, NY.
50. Peschel, N. and C. Helfrich-Förster, *Setting the clock-by nature: circadian rhythm in the fruitfly Drosophila melanogaster.* FEBS letters, 2011.
51. Hardin, P.E., *5 Molecular Genetic Analysis of Circadian Timekeeping in Drosophila.* Advances in Genetics, 2011. **74**: p. 141.
52. Glossop, N.R.J., L.C. Lyons, and P.E. Hardin, *Interlocked feedback loops within the Drosophila circadian oscillator.* Science, 1999. **286**(5440): p. 766.
53. Saez, L., P. Meyer, and M. Young. *A PER/TIM/DBT interval timer for Drosophila's circadian clock*. 2007. Cold Spring Harbor, NY.
54. Dunlap, J.C., *PHYSIOLOGY: Enhanced: Running a Clock Requires Quality Time Together.* Science's STKE, 2006. **311**(5758): p. 184.
55. Meyer, P., L. Saez, and M.W. Young, *PER-TIM interactions in living Drosophila cells: an interval timer for the circadian clock.* Science's STKE, 2006. **311**(5758): p. 226.
56. Lee, C., K. Bae, and I. Edery, *PER and TIM Inhibit the DNA Binding Activity of aDrosophila CLOCK-CYC/dBMAL1 Heterodimer without Disrupting Formation of the Heterodimer: a Basis for Circadian Transcription.* Molecular and cellular biology, 1999. **19**(8): p. 5316-5325.
57. Koh, K., X. Zheng, and A. Sehgal, *JETLAG resets the Drosophila circadian clock by promoting light-induced degradation of TIMELESS.* Science's STKE, 2006. **312**(5781): p. 1809.
58. Peschel, N., et al., *Light-dependent interactions between the Drosophila circadian clock factors cryptochrome, jetlag, and timeless.* Current biology, 2009. **19**(3): p. 241-247.
59. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-679.

60. Emery, P., et al., *Drosophila cryptochromes: a unique circadian-rhythm photoreceptor.* Nature, 2000. **404**(6777): p. 456-457.
61. Chaves, I., et al., *The cryptochromes: blue light photoreceptors in plants and animals.* Annual review of plant biology, 2011. **62**: p. 335-364.
62. Sancar, A., *Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors.* Chemical Reviews-Columbus, 2003. **103**(6): p. 2203-2238.
63. Sancar, A., *Photolyase and cryptochrome blue-light photoreceptors.* Advances in protein chemistry, 2004. **69**: p. 73-100.
64. Ceriani, M.F., et al., *Light-dependent sequestration of TIMELESS by CRYPTOCHROME.* Science, 1999. **285**(5427): p. 553-556.
65. Collins, B., et al., *Drosophila CRYPTOCHROME is a circadian transcriptional repressor.* Current biology, 2006. **16**(5): p. 441-449.
66. Chen, K.F., et al., *QUASIMODO, a Novel GPI-Anchored Zona Pellucida Protein Involved in Light Input to the Drosophila Circadian Clock.* Current biology, 2011.
67. Müller, M. and T. Carell, *Structural biology of DNA photolyases and cryptochromes.* Current opinion in structural biology, 2009. **19**(3): p. 277-285.
68. EKER, A., et al., *Direct DNA damage reversal: elegant solutions for nasty problems.* Cellular and molecular life sciences, 2009. **66**(6): p. 968-980.
69. Li, J., et al., *Dynamics and mechanism of repair of ultraviolet-induced (6-4) photoproduct by photolyase.* Nature, 2010. **466**(7308): p. 887-890.
70. Liu, Z., et al., *Electron Tunneling Pathways and Role of Adenine in Repair of Cyclobutane Pyrimidine Dimer by DNA Photolyase.* Journal of the American Chemical Society, 2012.
71. Thiagarajan, V., et al., *Kinetics of cyclobutane thymine dimer splitting by DNA photolyase directly monitored in the UV.* Proceedings of the National Academy of Sciences, 2011. **108**(23): p. 9402.
72. Liu, Z., et al., *Dynamics and mechanism of cyclobutane pyrimidine dimer repair by DNA photolyase.* Proceedings of the National Academy of Sciences, 2011. **108**(36): p. 14831-14836.

73. Brettel, K. and M. Byrdin, *DNA photolyase: Is the nonproductive back electron transfer really much slower than forward transfer?* Proceedings of the National Academy of Sciences of the United States of America, 2012.
74. Zhong, D., A. Sancar, and A. Stuchebrukhov, *Reply to Brettel and Byrdin: On the efficiency of DNA repair by photolyase.* Proceedings of the National Academy of Sciences, 2012.
75. Selby, C.P. and A. Sancar, *The Second Chromophore in Drosophila Photolyase/Cryptochrome Family Photoreceptors.* Biochemistry, 2011.
76. Geisselbrecht, Y., et al., *CryB from Rhodobacter sphaeroides: a unique class of cryptochromes with new cofactors.* EMBO reports, 2012. **13**(3): p. 223-229.
77. Li, Q.H. and H.Q. Yang, *Cryptochrome Signaling in Plants†.* Photochemistry and photobiology, 2007. **83**(1): p. 94-101.
78. Shalitin, D., et al., *Blue light–dependent in vivo and in vitro phosphorylation of Arabidopsis cryptochrome 1.* The Plant Cell Online, 2003. **15**(10): p. 2421-2429.
79. Burney, S., et al., *Conformational change induced by ATP binding correlates with enhanced biological function of Arabidopsis cryptochrome.* FEBS letters, 2009. **583**(9): p. 1427-1433.
80. Mohawk, J.A., C.B. Green, and J.S. Takahashi, *Central and Peripheral Circadian Clocks in Mammals.* Annual Review of Neuroscience, 2012(0).
81. Sancar, A., *Regulation of the mammalian circadian clock by cryptochrome.* Journal of Biological Chemistry, 2004. **279**(33): p. 34079-34082.
82. Lamia, K.A., et al., *AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation.* Science's STKE, 2009. **326**(5951): p. 437.
83. Lamia, K.A., et al., *Cryptochromes mediate rhythmic repression of the glucocorticoid receptor.* Nature, 2011. **480**(7378): p. 552-556.
84. Hoang, N., et al., *Human and Drosophila cryptochromes are light activated by flavin photoreduction in living cells.* PLoS biology, 2008. **6**(7): p. e160.
85. Foley, L.E., R.J. Gegear, and S.M. Reppert, *Human cryptochrome exhibits light-dependent magnetosensitivity.* Nature communications, 2011. **2**: p. 356.

86. Vieira, J., et al., *Human Cryptochrome-1 Confers Light Independent Biological Activity in Transgenic Drosophila Correlated with Flavin Radical Stability.* PloS one, 2012. **7**(3): p. e31867.
87. Stanewsky, R., et al., *The< i> cryb</i> Mutation Identifies Cryptochrome as a Circadian Photoreceptor in< i> Drosophila</i>.* Cell, 1998. **95**(5): p. 681-692.
88. Busza, A., et al., *Roles of the two Drosophila CRYPTOCHROME structural domains in circadian photoreception.* Science's STKE, 2004. **304**(5676): p. 1503.
89. Dissel, S., et al., *A constitutively active cryptochrome in Drosophila melanogaster.* Nature neuroscience, 2004. **7**(8): p. 834-840.
90. Öztürk, N., et al., *Animal Type 1 Cryptochromes.* Journal of Biological Chemistry, 2008. **283**(6): p. 3256-3263.
91. Liu, B., et al., *Searching for a photocycle of the cryptochrome photoreceptors.* Current opinion in plant biology, 2010. **13**(5): p. 578-586.
92. Ozturk, N., et al., *Reaction mechanism of Drosophila cryptochrome.* Proceedings of the National Academy of Sciences, 2011. **108**(2): p. 516-521.
93. Song, S.H., et al., *Formation and function of flavin anion radical in cryptochrome 1 blue-light photoreceptor of monarch butterfly.* Journal of Biological Chemistry, 2007. **282**(24): p. 17608.
94. Ozturk, N., et al., *Comparative photochemistry of animal type 1 and type 4 cryptochromes.* Biochemistry, 2009. **48**(36): p. 8585-8593.
95. Fogle, K.J., et al., *CRYPTOCHROME is a blue-light sensor that regulates neuronal firing rate.* Science's STKE, 2011. **331**(6023): p. 1409.
96. Krishnan, B., et al., *A new role for cryptochrome in a Drosophila circadian oscillator.* Nature, 2001. **411**(6835): p. 313-317.
97. Gegear, R.J., et al., *Cryptochrome mediates light-dependent magnetosensitivity in Drosophila.* Nature, 2008. **454**(7207): p. 1014-1018.
98. Yoshii, T., M. Ahmad, and C. Helfrich-Förster, *Cryptochrome mediates light-dependent magnetosensitivity of Drosophila's circadian clock.* PLoS biology, 2009. **7**(4): p. e1000086.
99. Gegear, R.J., et al., *Animal cryptochromes mediate magnetoreception by an unconventional photochemical mechanism.* Nature, 2010. **463**(7282): p. 804-807.

100. Zoltowski, B.D., et al., *Structure of full-length Drosophila cryptochrome.* Nature, 2011. **480**(7377): p. 396-399.

Chapter 2

STRUCTURE OF LIGHT-ACTIVATED LOV PROTEIN DIMER THAT REGULATES TRANSCRIPTION*

2.1 Abstract

Light oxygen or voltage (LOV) domains are widely represented signaling modules in bacteria, archea, protists, plants and fungi. The *Neurospora crassa* LOV protein VIVID (VVD) allows adaptation to constant or increasing light levels and proper entrainment of circadian rhythms. The crystal structure of the fully light-adapted VVD dimer reveals the mechanism by which light driven conformational change alters oligomeric state. Photo-induced formation of a cysteinyl-flavin adduct generates a new hydrogen bond network that releases the N-terminus from the protein core and restructures an acceptor pocket for its binding on the opposite subunit. Substitution of residues key to the monomer/dimer switch have profound effects on light adaptation in *Neurospora*. The VVD dimerization mechanism provides the molecular details for how a large family of photoreceptors converts light responses to alterations in protein interactions.

* The material in this chapter is reprinted with permission from A. T. Vaidya, C.-H. Chen, J. C. Dunlap, J. J. Loros, B. R. Crane, Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription. Science Signaling 4, ra50 (2011). Copyright (2011) AAAS.

2.2 Introduction

Per-ARNT-Sim (PAS) domains (~120 residues) respond to environmental change through chemical reactions of bound cofactors (e.g., flavin, heme, Fe-S cluster, etc.), by the binding of exogenous ligands, and/or by alteration of oligomeric state [1-3]. The light oxygen or voltage (LOV) family of PAS proteins binds flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN) and senses blue-light through photochemically driven formation of a cysteinyl-flavin adduct [4-6]. Two such LOV domain-containing proteins manifest light-induced resetting and phasing of the *Neurospora crassa* circadian clock [7-10] and are the founding members of a principal class of photoreceptors conserved among all fungi [11]. The first, White collar-1 (WC-1), is a large multi-domain transcription factor [12, 13] that along with White Collar-2 (WC-2) comprises the White Collar Complex (WCC), the primary photoreceptor for the *Neurospora* circadian system [8, 9, 14, 15]. In response to blue light, the WCC induces transcription of most light-responsive genes [16], including many clock-controlled genes (CCGs) as well as the gene encoding the central circadian oscillator protein FREQUENCY (FRQ) [8, 9, 11-13]. The second light sensor, VIVID (VVD), strongly induced by the WCC in response to light, antagonizes the action of the WCC by interacting directly with it, and in doing so adapts the organism to constant and increasing light levels [7, 17-21]. VVD is almost entirely composed of a LOV domain that has high homology to the LOV domain of WC-1 [7, 18]. Thus, in WC-1 and VVD, two similar light sensory modules exhibit very different biological activities in the same signaling circuit. Currently there are several examples of LOV protein crystal structures; however, none are characterized in both “on” and “off” conformations that have been

verified as biologically relevant signaling states [22-24]. Elegant solution NMR studies on the *Avena sativa* phototropin1 LOV2 domain reveal how a C-terminal α-helix displaces from the core LOV domain upon light excitation [25], but a detailed structural description for how the remodeled LOV domain affects enzymatic activity in the full-length protein is currently lacking for this system [26] and others [27]. Light induces VVD homodimerization, thereby providing a mechanism for how photon absorption can be converted to the remodeling of protein interactions important for signaling [28].

Previously, we reported a light-state structure of VVD that was determined by exposing dark state crystals to white light [22]. These crystals contain a majority population of the cysteinyl-flavin adduct. The structure shows a set of structural rearrangements that result from protonation of flavin atom N5 and propagate through a conserved hinge to the N-terminal cap (N-cap) region of the PAS domain. One of the most variable regions in PAS domains, the N-cap, packs against the β-sheet of the PAS core and in VVD is composed of an alpha helix (aα), a short β-strand (bβ) and an N-terminal “latch” (residues 37-44) that wraps around the domain, toward the surface-exposed adenosine moiety of FAD. The largest structural changes in the light-exposed crystals involve a ~4 Å shift of the bβ strand [22]. However, in solution, adduct formation drives greater conformational change at the N-terminus and subsequent dimerization [22, 28]. The N-cap connects to the first β-strand of the PAS core (Aβ) through a hinge region that contains Cys71. Substitution of the hinge residue Cys71Ser, which changes from a buried to an exposed position in the irradiated crystals, prevents conformational change at the N-terminus, inhibits dimerization *in vitro*, and curtails VVD function *in vivo* [20-22, 24, 29]. In contrast, the Cys71Val variant, predisposes the hinge

to the light-state conformation, releases the N-terminus and generates some dimer, even in the dark [24, 28, 30]. Thus, the light-exposed dark-state crystal represents an intermediate conversion to the fully light-adapted state: conformational changes within the protein have taken place, but the crystal lattice has prevented large-amplitude re-arrangements and subunit association. Herein, we describe the crystallographic structure of the VVD fully activated light-state dimer. The structure contrasts with the monomeric form of the VVD dark-state and represents an activated LOV protein in complex with its immediate target, in this case, another subunit of the same light sensor. The design of VVD variants based on the structure, perturb *Neurospora* photosensory transduction in predictable ways.

2.3 Results

2.3.1 Trapping the light-state using a VVD variant

Although the adduct lifetime of VVD (~5 h) is long compared to many LOV proteins, it is not long enough to crystallize the light-state dimer. We thus turned to a Met135Ile:Met165Ile variant (VVD-II) that was designed to increase adduct lifetime by stabilizing the reduced form of the flavin ring [31]. Adduct lifetime increases 10 fold in VVD-II and the recombinant protein can be purified from *E. coli* in a partial neutral semiquinone state, which is indicative of a higher flavin redox potential (Figure 2.1). After white light exposure, the protein grew crystals overnight that were clear and diffracted to 2.7 Å resolution. To limit further reduction and cleavage of the adduct state in the synchrotron X-ray beam, diffraction data were collected from 10 different positions on a single crystal under conditions that limited radiation exposure.

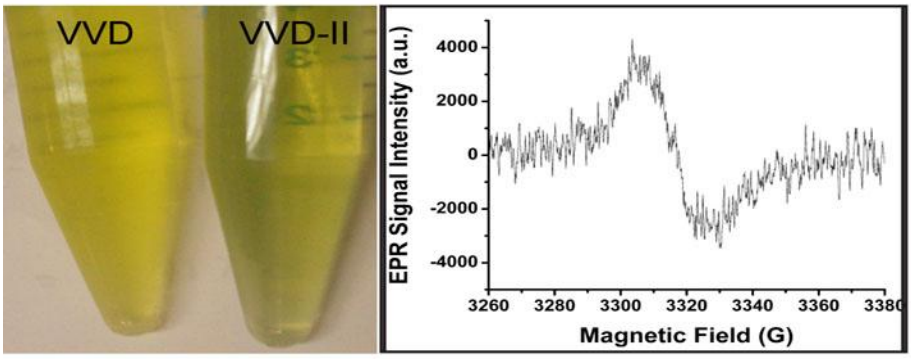


Figure 2.1: Low redox potential of VVD-II. VVD-II (Met135Ile:Met165Ile variant) purifies from *E. coli* in a partial neutral semiquinone state, which is indicated by the green colour of the protein compared to the yellow colour of the native protein (left) and the Electron Paramagnetic Resonance (EPR) spectrum of VVD-II (9 GHz cw EPR spectrometer, right).

Table 2.1: Data Collection and Phasing Statistics

Wavelength (Å)	0.97918
Space group	$P2_1$
Cell Parameters (Å)	a = 37.59, b = 77.55, c = 54.78, β= 99.73
Resolution (Å)	30 - 2.75 (2.85 - 2.75)
No. of observations	31885
No. of unique reflections	8227
Completeness (%)	98.6 (97.9)
R_{sym} [a]	0.134 (0.485)
I/σ(I)	10.3 (2.0)

Refinement statistics

Resolution range (Å)	30-2.75 (2.85 - 2.75)
R factor, %	22.33 (34.67)

R_{free}, %	28.76 (39.32)
Wilson B factor ($Å^2$)	42.1
Molecules / Asym unit	2
Residues	2 X 149
Atoms	
Protein	2346
Solvent	207
Cofactor (FAD)	106
Mean B-values ($Å^2$)	
Overall	41.7
Main chain	40.8
Side chain	42.7
Solvent	33.0
Rmsd from ideal geometry	
Bonds (Å)	0.017
Angles (°)	2.14
Ramachandran plot, %	
Most favored	83.6
Additionally allowed	15.6
Generously allowed	0.8
Disallowed	0.0
Missing residues	None

Data for outermost resolution shell are given in parenthesis.

[a] $R_{sym} = \Sigma\Sigma_j |I_j - \langle I \rangle| / \Sigma\Sigma_j I_j$

2.3.2 Light-induced conformational changes and the mechanism of light-signaling

As expected, light-activated VVD-II crystallizes as a dimer in a new crystal lattice, with one dimer per asymmetric unit (Table 2.1). The fully formed adduct has strong electron density between the Cys108 thiol group and flavin C4a and noticeable puckering of the isoalloxazine ring at C4a (Figure 2.2A). Overlays of the structures for the light-state dimer (LSD), the intermediate-state monomer (ISM) and the dark-state monomer (DSM) reveal that the LSD not only undergoes many of the internal structural rearrangements found in the ISM structure, but also shows much larger changes in the N-cap region where the N-terminal latch has been released from the subunit core (Figure 2.2B, C, D). In the ISM structure (at 1.7 Å resolution), protonation of flavin N5 causes Gln182 to flip and alter hydrogen bonding interactions with the backbone of Ala72 and the hinge region to the N-cap. In the LSD, at 2.7 Å resolution, the Gln182 side chain amide orientation cannot be assigned unambiguously based on electron density, but changes in its position and that of the Ala72 carbonyl compared to the DSM reflect the expected flip (Figure 2.2C).

The hinge region (residues 65 to 72) plays a major role in the light activated conformational changes in VVD. In both the ISM, and the LSD, formation of the new hydrogen bond between the Ala72 carbonyl and Gln182 couples to rearrangement of the hinge region. In both structures, Cys71 is released from its buried position in the DSM and no longer hydrogen bonds to the backbone of Asp68 (Figure 2.2C). In the LSD, however, the Cys71 thiol group orients toward the Ala72 carbonyl and not the Asp68 peptide nitrogen, as found in the ISM structure (Figure 2.2C). Similar to the ISM

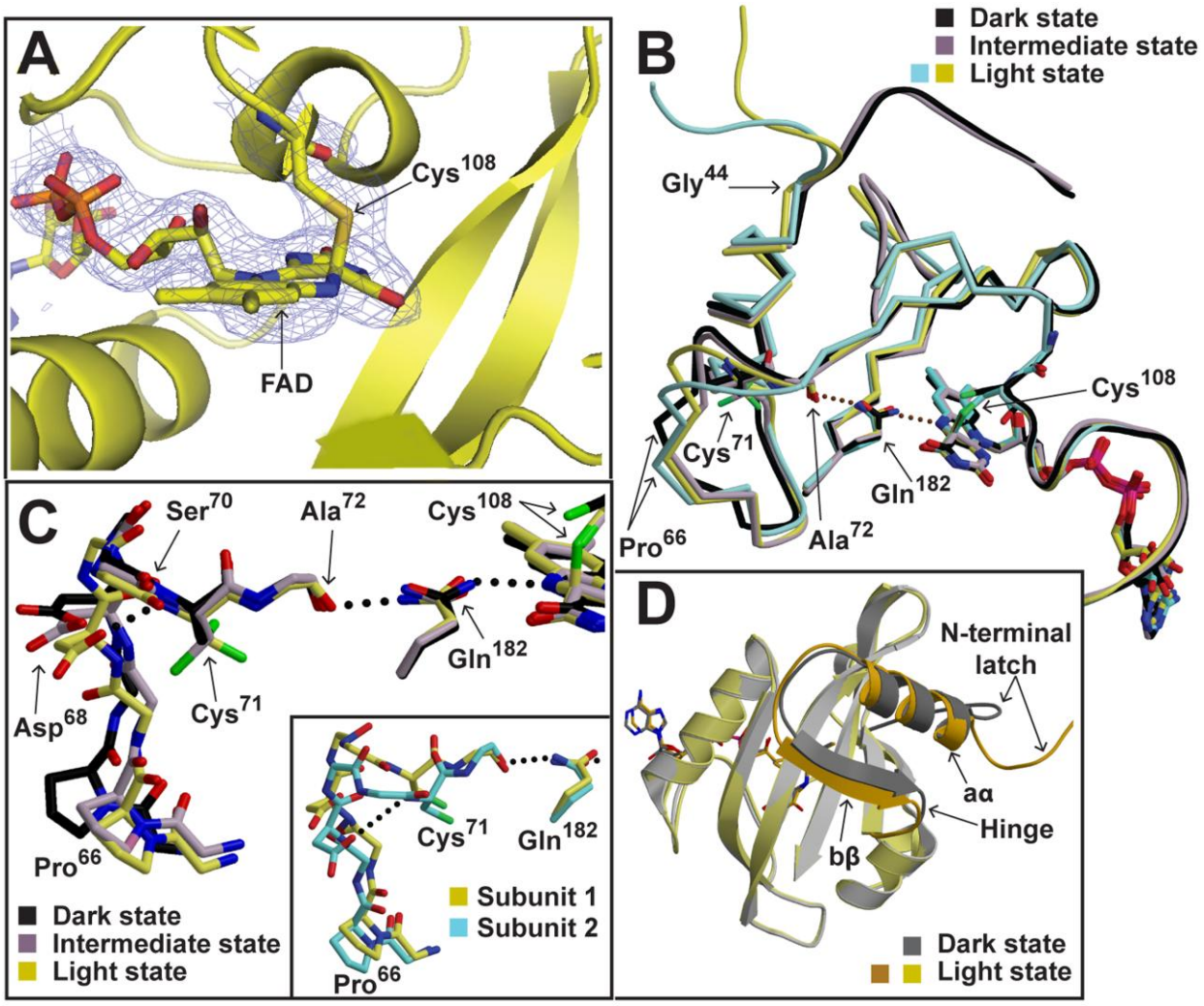


Figure 2.2: Light-state structure of VVD and the mechanism of light-induced signal transduction. (A) Electron density corresponding to the light induced adduct between the sulfur of Cys108 and the C4a carbon of the FAD (F_o-F_c omitmaps, blue contures at 2.5σ). (B, C) Blue-light induced adduct formation reduces the C4a-N5 bond of the flavin. In response to the newly protonated N5, Gln182 flips and forms two new hydrogen bonds – one with the N5 proton and the other with Ala72 carbonyl oxygen. This shifts the Ala72 and breaks the hydrogen bond between Cys71 thiol and Asp68. A new hydrogen bond forms between Asp68 and the main chain nitrogen of Ser70. A slight shift of Ala72 leads to systematic changes in the hydrogen bonding pattern in the hinge region, resulting in a 3.5 Å shift of Pro66. Shifts of aα, destabilize the N-terminal latch against the protein core, and leads to its release, which is further facilitated by the flexibility of Gly44. Inset: Changes in the hinge loop shown in yellow are sufficient to respond to adduct formation; however, in the other subunit of the LSD (cyan) the loop moves to a greater extent. (D) In the light state structure, the N-cap (represented in darker shades of grey and yellow) restructure to release the N-terminal latch while the other regions of the protein do not change appreciably compared to the dark-state structure.

structure, bβ shifts toward the PAS β-sheet with Pro66 moving by ~3.5 Å. Although the hinge loop containing Cys71 has a similar conformation in both subunits of the LSD, it was displaced further over bβ in one subunit than the other (Figure 2.2C, inset). The N-terminal shifts further over bβ in one subunit than the other (Figure 2.2C, inset). The N-terminal helix shifts along its axis to a greater extent (~1.4 Å) than that observed in the ISM, where it moves very little (Figure 2.2D). The helical shifts are within the range that do not require repacking of side chains [32] and appear to couple directly to repositioning of bβ and the Cys71-containing hinge loop, against which aα packs. Movements of aα and bβ destabilize contacts of the N-terminal latch, which assumes an entirely new conformation compared to that of the DSM and the ISM by the backbone rotation of Gly44 (Figure 2.2B, D). We cannot rule out the possibility that other long-range effects such as changes in dipole and charge distributions associated with the adduct state contribute to latch release, but given the distance of separation and requirement of residues in the connecting loop for propagating the signal, the influence of through-space factors is likely less than conformational changes propagated through bβ and aα.

2.3.3 Light-induced dimerization and its implications

The N-terminal latch, freed from its association with the PAS core, docks into the hinge region of the adjacent subunit (Fig 2.3). Thus, by restructuring of the hinge, photo-induced adduct formation not only facilitates release of the N-terminus on one subunit, but also provides a recognition pocket to accept the latch from the opposite subunit. Indeed, previous time-resolved SAXS data of VVD variants with altered dimerization

properties indicate the release of the N-terminus is necessary but not sufficient for dimer formation [30]. At the center of the extensive dimer interface (1342 $Å^2$ surface area per subunit; Predicted free energy of formation (ΔG_i) = -26.8 kcal mol^{-1}; Hydrophobic specificity factor (ΔG_p) = 0.009 [33]). Tyr40 inserts into the hinge pocket created by bβ and the C-terminal β-strand, Eβ and hydrogen bonds with both the side-chain and main-chain of Eβ and the Cys71 thiol (Figure 2.3). Although the hinge loops differ in position by ~4 Å in the two LSD subunits, their conformations and interactions with Tyr40 are similar. The N-terminal helix aα also contributes to the interface through the symmetric contacts of an exposed hydrophobic face involving Met48 and Ile52 (Figure 2.3). The molecular envelope of the transient dimer species observed in time-resolved SAXS experiments [30] agrees well with the overall dimensions of the LSD structure and much better than those of non-physiological VVD dimers found in crystal lattices (Figure 2.4).

The behavior of many previously studied VVD variants are explained by the LSD structure, including the requirement of the latch residues (37-44) for high-affinity dimerization (20). Truncations beyond Gly44, will not dimerize under any conditions [28]. Substitutions within the hinge are expected to have complex effects on dimerization depending on their propensity to release the latch and/or restructure the recognition pocket. Cys71Ser prevents dimerization by short-circuiting latch release and preventing latch binding [22], while Cys71Val produces an extended monomer with the latch released and some dimerization in the dark state [28]. Substitution of contacts along the interaction surface of aα can prevent dimerization (e.g. Met55Arg) [22] whereas changes on the internal surface of the helix (Tyr50Trp) promote light-induced dimerization by presumably destabilizing aα packing and facilitating latch release [28].

Chemical modifications at Cys183, which lies at the end of Eβ and contacts the N-terminal latch, prevents dimerization [22]. Light-dependent disulfide cross-linking of Glu171Cys is much better explained by the LSD structure than a crystallographic dimer formed by the DSM [28].

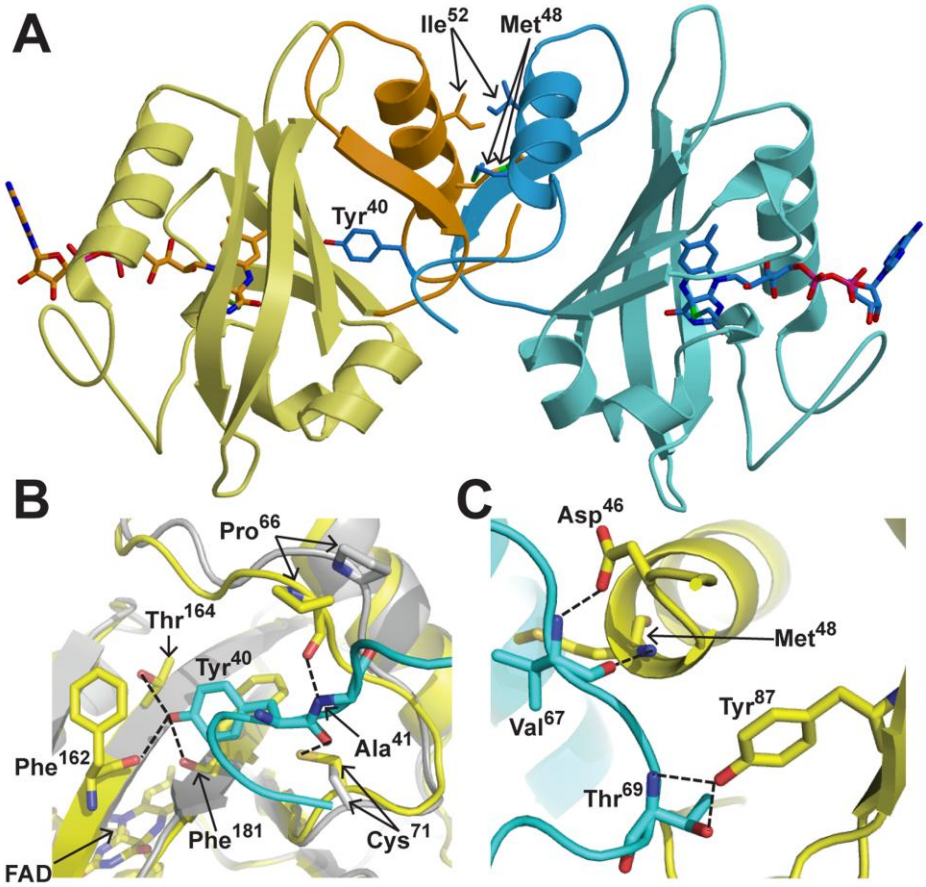


Figure 2.3: The light state dimer of VVD. (A) The protein crystallizes as a dimer with the N-cap (represented in darker shades of yellow and cyan) composing the dimer interface. Met48 and Ile52 in aα make important hydrophobic interactions within the interface. (B) Interactions within the dimer interface: Tyr40 (cyan) hydrogen bonds with the side chain of Thr164 and the main chains of Phe162 and Phe181 of the other subunit. The main chain carbonyl of Pro66, which shifts by 3.5Å, hydrogen bonds with the main chain nitrogen of Ala41 of the other subunit. Cys71, which flips in the light state, hydrogen bonds with the main chain of Tyr40 in the opposite subunit. (C) Also, in the dimer interface Tyr87 hydrogen bonds with the side chain and the backbone of Thr69, whereas Met48 and Asp46 hydrogen bond with the backbone of Val67.

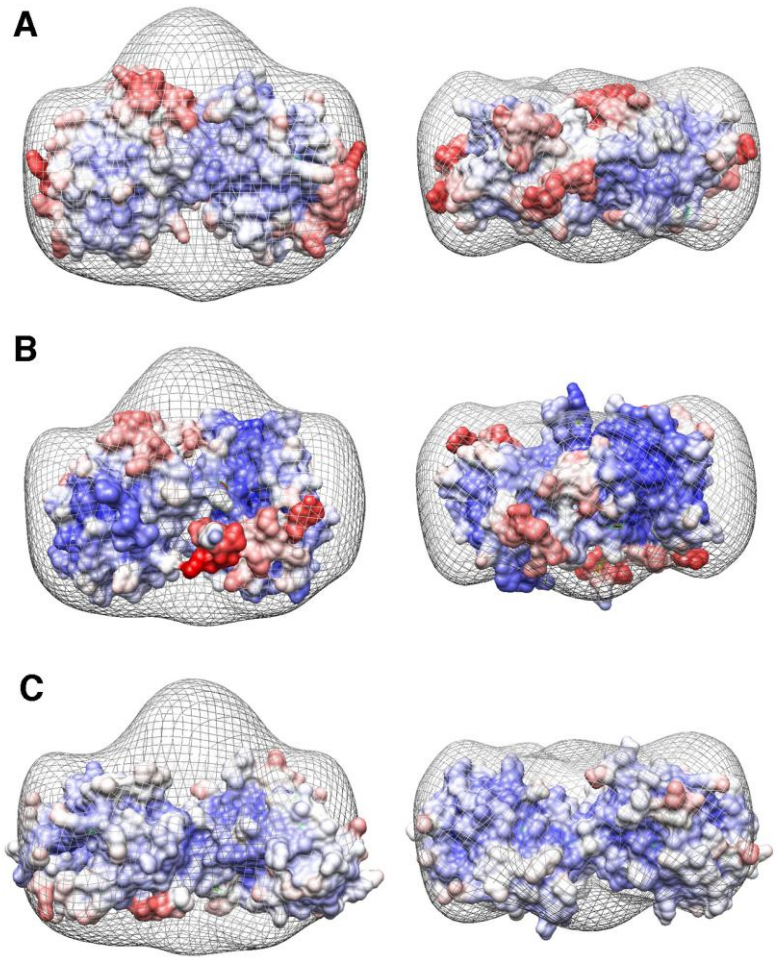


Figure 2.4: Comparison of VVD LSD structure to time-resolved SAXS data. Superposition of the molecular envelope determined from time-resolved SAXS of the VVD light-state and the solvent accessible surface of several VVD crystallographic dimers colored by relative thermal B-factor (blue – low; red - high). View on the right is down from the top as shown on the left. The dimensions of the SAXS envelope generally match that of the VVD LSD (**A**), with the exception of a protrusion of SAXS density along the dimer twofold axis. However, this small discrepancy likely reflects dynamics in this part of the molecule; extended, mobile regions of the VVD structure (denoted by the red surface), point toward the dimer twofold axis at the top of the envelope (Note that although the position of the structure within the envelope is the best fit, it is not unambiguous). The procedure that calculates the SAXS envelope (See J. S. Lamb *et al., J. Mol. Biol.* **393** (2009)) does not account for conformational heterogeneity and rather assumes a constant protein density and sharp boundary between solvent and protein, which in the more mobile regions of the protein, is not a fully representative model. Other crystallographically characterized VVD dimers do not fit the SAXS envelope nearly as well as the LDS structure. Panel **B** shows the alternative non-physiological dimer present in the LSD crystal lattice and panel **C** shows the VVD dimer found in dark-state crystals (PDB Code: 2PD7).

2.3.4 Confirming the light-induced dimeric structure in solution

To further confirm that the LSD forms in solution and create variants for use *in vivo*, we examined other residue substitutions that disrupt the hydrogen bonds at the dimer interface. As expected, mutations at Tyr40 drastically reduced light-triggered association, with Tyr40Lys completely abrogating dimerization (Figure 2.5). The

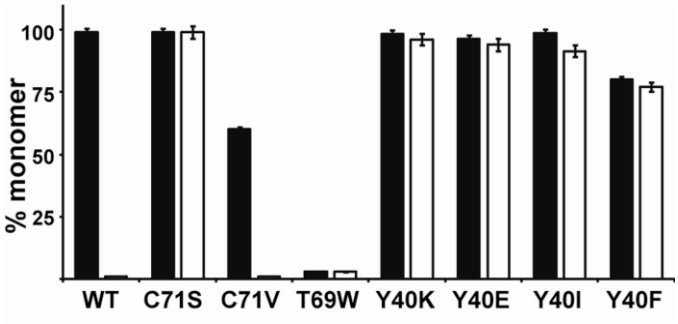


Figure 2.5: Ability of VVD variants to undergo dimerization. The % monomer is calculated by areas of respective light and dark state elution peaks on size exclusion chromatography [28]. The Y40K, Y40E, Y40I and Y40F variants disrupt important hydrogen bonds at the interface and lose the ability to dimerize in the presence of light. The T69W variant is a dimer in both the dark and the light, presumably because it facilitates intersubunit contacts that overcome light-promoted conformational switching. (* from reference 22, ** from reference 26)

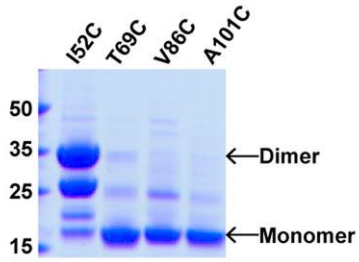


Figure 2.6: Targeting cross-linking of the VVD LSD. The Ile52Cys spontaneously cross-links in the light when purified from *E. coli*, whereas other positions (Val86 and Ala101) removed from the dimer interface, do not crosslink. Band at 25 kDa corresponds to proteolytically degraded VVD dimer. The Thr69Cys substitution also places a thiol at the dimer interface but subunits do not crosslink readily. This may be due to the non-ideal orientation of the cysteine thiols for disulfide formation.

Tyr40Phe variant reduces the LSD population, but also promotes some dark-state dimer formation, which suggests that it decouples subunit association from the conformational signals initiated by the adduct state (Figure 2.5). In the extreme, Thr69Trp dimerizes in both the dark and the light, presumably because it facilitates inter-subunit contacts that overcome light promoted conformational switching (Figure 2.5). We also tested interface positions for enhanced cross-linking and found that Ile52Cys spontaneously cross-links in the light when purified from *E. coli*, whereas other positions removed from the dimer interface, do not (Figure 2.6).

2.3.5 Role of the dimer in photoadaptation of Neurospora

Armed with VVD variants that both enhance and abrogate light-induced dimer formation without perturbing VVD photochemistry, we tested whether such alternations could affect *Neurospora* light sensing when expressed from a corresponding allele in a Δ*vvd* strain (Figure 2.7). *(This work was done by C.-H. Chen, J. C. Dunlap and J. C. Loros at the Dartmouth Medical School)* Consistent with previous studies [21] , the Cys71Ser and Cys108Ala show a partial loss-of-function phenotype as assayed by coloration in constant light (Figure 2.8A). Carotenoid production, which turns cells orange, is activated by WC-1, but inhibited by VVD. The Cys71Val, Thr69Trp and Ile52Cys mutant alleles appear to have a normal coloration phenotype (Figure 2.8A). The Tyr40Lys mutant allele, which is incapable of light induced dimerization, displayed high levels of carotenoid accumulation similar to a Δ*vvd* strain (Figure2.8A, B, Figure S4A). We observed no difference when strains were kept in constant darkness (DD, Figure 2.8A), indicating that the coloration phenotypes are light-dependent.

To quantify the photoadaptation defects at the molecular level, we compared the level of *al-3* repression among various *vvd* mutant strains held for 60 m in constant light (LL60) by RT-QPCR analysis. The expression of *al-3* is strongly induced by light (Figure 2.7B) and unambiguously displays photoadaptation in constant light in that, within an

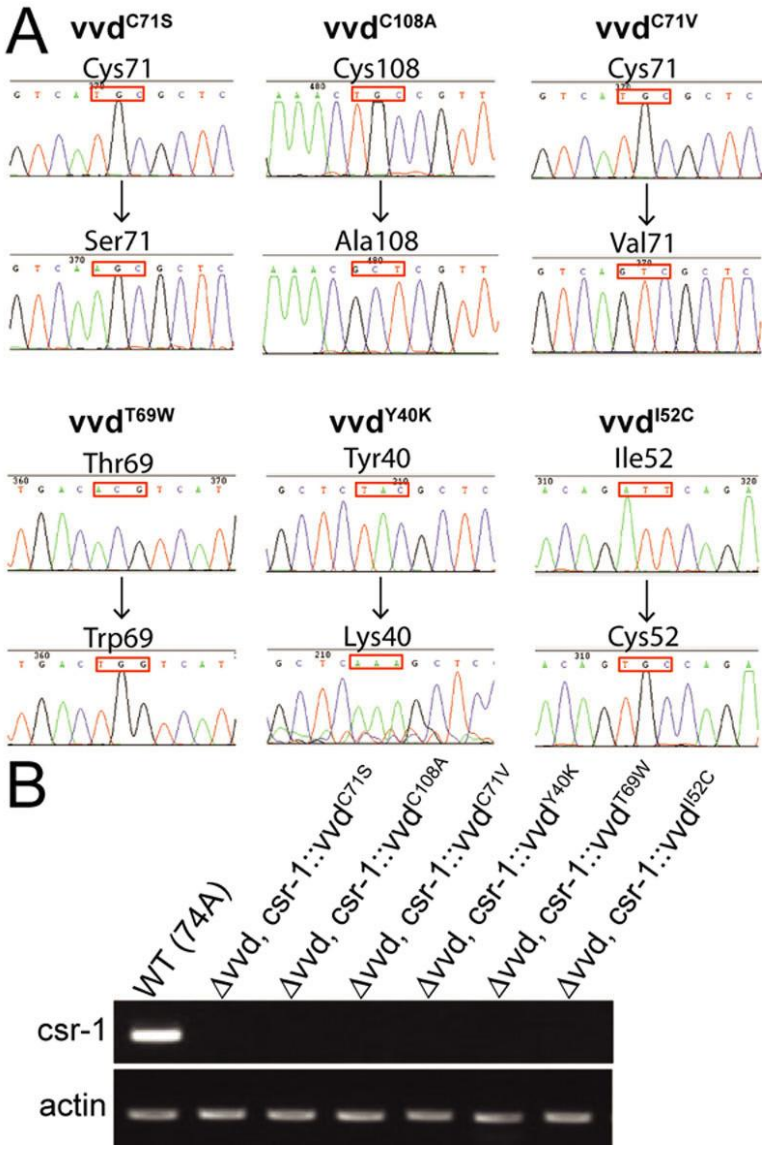


Figure 2.7: The *csr-1* knock-in strains used in this study are homokaryon. (A) Sequencing data of various *vvd* mutant alleles. (B) PCR analysis for the absence of the *csr-1* locus in the genome for each knock-in strain.

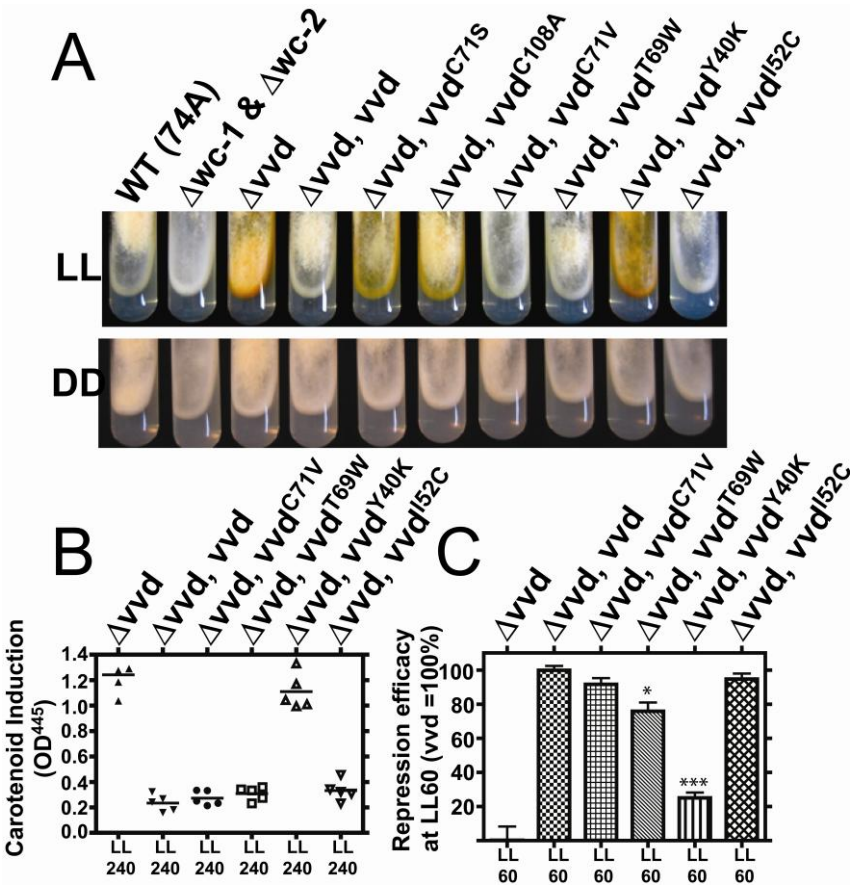


Figure 2.8: Dimerization of VVD is essential for the biological functions of VVD *in vivo.* (A) Carotenoid accumulation as a measure of photoadaptation defects in various *vvd* mutants. LL indicates strains were exposed to constant white light stimulus with photon flux of 20 µmol/m^2/sec for four days; DD indicates strains were kept in constant darkness. WT (74A) represents a wild-type strain without photoadaptation defects. *Δwc-1 & Δwc-2* represents a "blind" strain. (B) Photoadaptation defects quantified by the amount of carotenoid extracted from mycelia at LL240 (n=5, biological replicates). Horizontal lines denote means. (C) Repression efficacy of various *vvd* mutant alleles as determined by the repression of *al-3* expression at LL60 with RT-QPCR analysis (n=3, mean values ± standard error). Asterisks indicate statistical significance when compared to the wild-type allele of *vvd* as determined by *unpaired t-test*, ***p < 0.001, *p < 0.05.

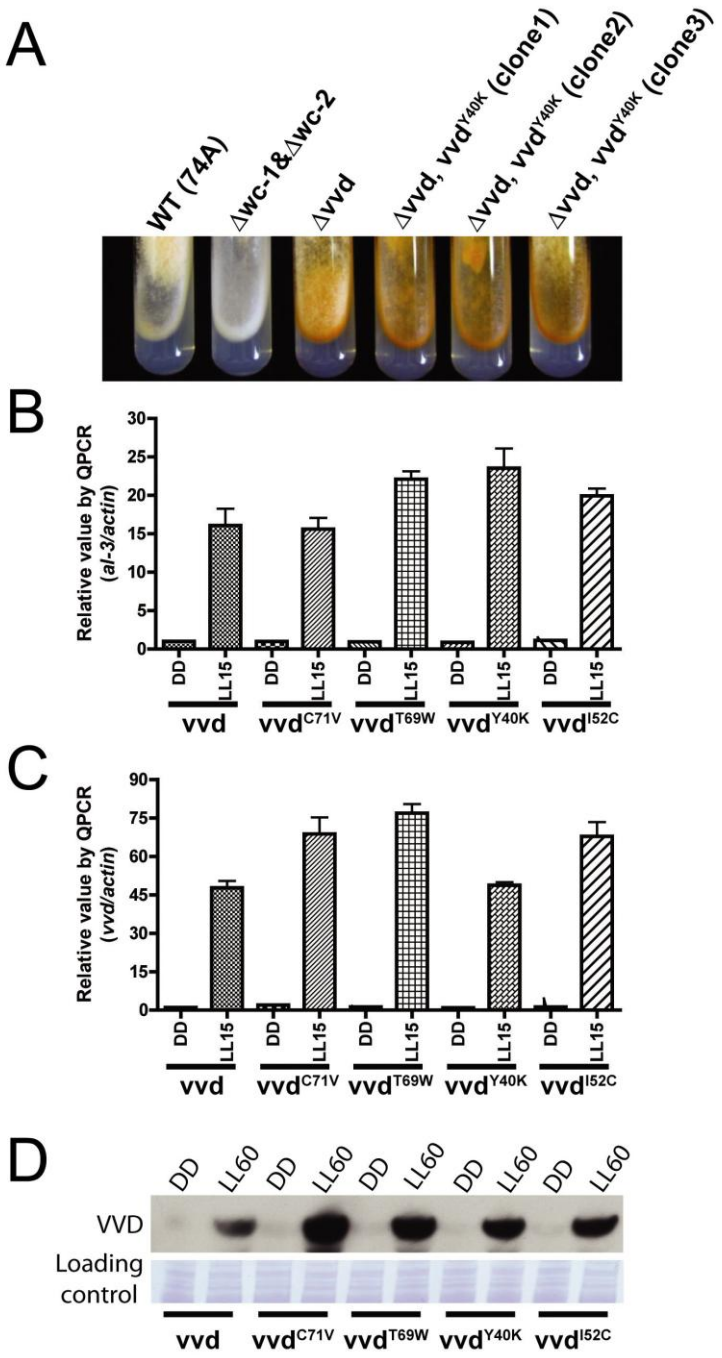


Figure 2.9: Light induction of *vvd* RNA and VVD protein in strains with various *vvd* mutant alleles. (A) Independent primary transformants of *vvd*^{Y40K} mutant allele (clones 1, 2, and 3) display a consistent phenotype of photoadaptation defects. (B) Light induction of *al-3* and (C) *vvd* RNA at LL15 in strains with various *vvd* mutant alleles determined by RT-QPCR analysis (n =3, mean values ± standard error). (D) Light induction of VVD protein at LL60 in strains with various *vvd* mutant alleles.

hour, levels of *al-3* expression were reduced to approximately the pre-induction level [21]. Repression efficacy (Figure 2.8C) was calculated based on the level of *al-3* repression between strains with either no *vvd* expression (0%, no repression) or a *vvd* wild-type allele (100%, maximal repression). Consistent with the coloration phenotype and high levels of carotenoid accumulation, the Tyr40Lys mutant allele displayed the most severe defect in repressing light responses (Figure 2.8C), retaining merely 25% of the wild-type VVD activity. Unexpectedly, the Thr69Trp mutant allele showed a slight but significant reduction in activity. Meanwhile, the extent of early light activation of *al-3* and *vvd* RNA at LL15 were more or less unaffected in various *vvd* mutants (Figure 2.9B and 2.9C), and VVD proteins were induced to comparable levels *in vivo* at LL60 (Figure 2.9). These data suggests that the photoadaptation defects are the result of mutations that interfere with the ability of VVD to undergo the crystallographically observed light-induced conformational changes that lead to dimerization.

2.4 Discussion

Residues important for the light driven dimerization mechanism of VVD are conserved by a large number of fungal LOV domain containing light sensors such as ENVOY, WC-1 and other GATA-type transcription factors (Table 2.2). In particular, the WC-1 LOV domain contains critical residues necessary for a VVD-like light-induced dimerization mechanism (Figure 2.10, Table 2.2). Light activation of WC-1 is also associated with an increase in oligomeric state [13, 19, 34] and swapping the LOV domain core of VVD (residues 71-186) for that of WC-1 maintains several primary light responses in *Neurospora* [18]. The current model for the inhibition of the WCC by VVD

requires direct interaction and involves competitive heterodimerization between the VVD LOV and the WC-1 LOV [19-21]. Sequence conservation (Figure 2.10) strongly suggests that the VVD:WC-1 heterodimer will most likely resemble the VVD light state dimer structure determined here. In support of this notion, the Thr69Trp mutant, which causes a constitutive VVD dimer is solution (Figure 2.5), shows attenuated inhibition of the WCC in cells (Figure 2.8), perhaps because the variant subunits do not exchange effectively with the WC-1 LOV domain.

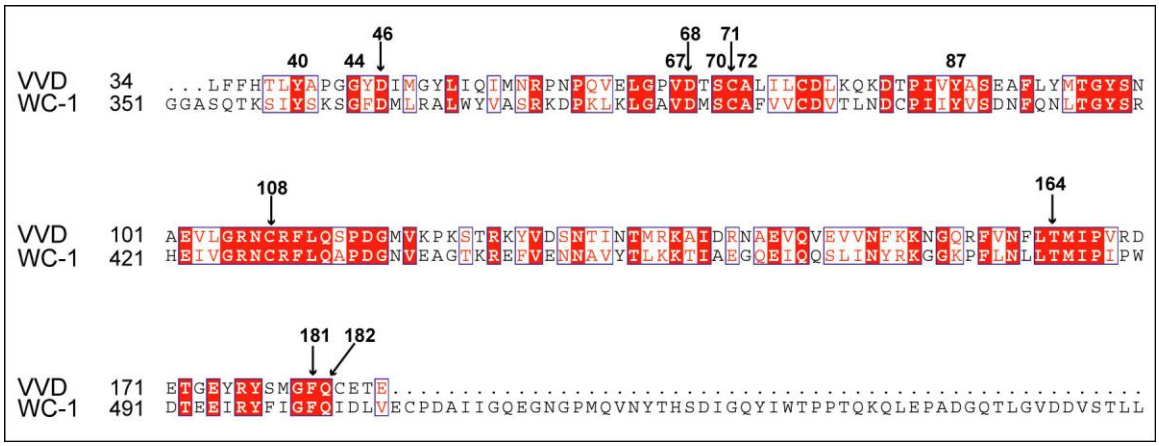


Figure 2.10: Sequence similarities between VVD and WC-1. Alignment of VVD and WC-1 protein sequences shows the conservation of 65% of the residues involved in light induced signaling and dimerization of VVD (Table S2). The residues involved in VVD light-sensing, signaling and dimerization and conserved in WC-1 are denoted with VVD numbering.

Table 2.2: Conservation of light-signaling and dimerization in the fungal LOV domains. (Left) The output of 170 protein sequences from Pipealign was classified into 10 groups. Ascomycota, Basidiomycota and Zygomycota are three phlya of the kingdom Fungi. UC stands for uncharacterized proteins. (Right) The residues involved in light-sensing (*sense*), conformational signaling (*signal*) and dimerization (*dimer*) in VVD and the corresponding residues across the ten groups are denoted as: ☑ - The residue is conserved in that group; ☐ - The corresponding residue is not conserved in the group; [-] - The corresponding residue is not present in the group; [X/Y] - The residues X and Y are conserved in the group instead of that found in VVD. Residues in green are a part of the N-terminal latch, residues in blue are in aα, those in magenta are in the hinge of VVD and those in red are conserved in all the groups.

Group (Number of sequences)	Proteins Representative primary accession numbers Classification of the sequences (number of sequences)
1 (7)	**VVD and its variants from *Neurospora crassa*** Q9C3Y6, Q1K5Y8 VVD (1) + VVD variants (4) + Predicted (2)
2 (22)	**GATA factors and predicted proteins from Ascomycota** Q7ZA36, A1D8C9, C8VHD4 GATA factor (11) + Predicted (8) + WC-1 (2)
3 (9)	**Putative WC-1 and WC-1 from Zygomycota** B5M9P6, Q14SS7, B8YIE4 WC-1 (2) + Predicted (7)
4 (15)	**Putative WC-1, VVD and ENVOY from Basidiomycota** A4RDB4, Q6DMM2, B8P2A7 WC-1 (3) + VVD (1) + ENVOY (1) + Predicted (7) + UC (3)
5 (21)	**WC-1 from Ascomycota** B2VXI8, C6K2I9, D2Y4Q8 WC-1 (19) + Predicted (1) + UC (1)
6 (16)	**Putative WC-1 and WC-1 from Ascomycota** A7F3V7, Q5PXG9, D1Z8I0 WC-1 (4) + Predicted (7) + UC (5)
7 (20)	**Phototropin and uncharacterized proteins from plants** C5YK68, Q5DW42, A7LI54 Phototropins (12) + Predicted (2) + UC (6)
8 (11)	**Phototropins from plants** B9T0R3, D7LMS0, Q8H934 Phototropins (7) + Predicted (3) + UC (1)
9 (12)	**Putative phototropins from plants** C5YPQ1, B8BLR3, O48547 Phototropins (4) + Predicted (5) + UC (3)
10 (37)	**Histidine kinase and sensor proteins** D5RKE3, A5EAE7, B7KW67 Histidine kinase (18) + Predicted (15) + UC (1)

	Group →	1	2	3	4	5	6	7	8	9	10
Dimer	Y40	✓	✓	✓	✓	✓	✓	E	D/E	E	
Dimer	A41	✓	S	S/A	S	S	S	R	R/S	R	
Signal	G44	✓	✓	✓	✓	✓	✓	S	S/D/E	S	
Dimer	D46	✓	✓	✓	✓	✓	✓	D/E	✓	E/D	
Dimer	M48	✓	M/A	I/L/V	L	L	L		K	E	
Dimer	I52	✓	A	S		R	Y/L/R	R	K/R	K	
Signal Dimer	P66	✓	A/P	✓	✓	S	A	R	R	R	V
Signal	V67	✓	✓	I/V	I	✓	✓	I	I	I	E/R
Signal Dimer	D68	✓	✓	✓	✓	✓	✓	E	E	E	T/A
Signal Dimer	T69	✓	L/M	L/M		L	M	K	K	K	✓
Signal	S70	✓	✓	✓	✓	✓	✓	N	N	N	R
Signal Dimer	C71	✓	✓	✓	C/V	✓	✓	–	–	–	M
Signal	A72	✓	✓	S	S/A	✓	✓	–	–	–	P/A
Dimer	V86	✓	V/I	V/I	V/I	✓	I	I	I	I	I/V
Dimer	Y87	✓	✓	✓	✓	✓	✓	F	F	F	F/Y
Sense	C108	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dimer	F162	✓	L	L	L	L	L	L	L	L	A
Dimer	T164	✓	✓	✓	✓	✓	✓	H	H	H	F/Y
Dimer	F181	✓	✓	L	✓	✓	✓	V	V	V	S
Signal	Q182	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

In other LOV-containing proteins, such as the plant and algal phototropin kinases, and bacterial sensors such as YtvA, the residues corresponding to Cys108 and Gln182 are strictly conserved, whereas those in the hinge region and N-cap differ (Table 2.2). The light-excited intermediate structure of *Avena sativa* phototropin1 LOV2 (*As* phot1 LOV2) [24] reveals that the conserved Gln side chain also flips in response to Cys-flavin adduct formation. As in VVD, the Gln rearrangement alters hydrogen bonding to peripheral elements, in this case the LOV2 Asn414, which, like Ala72 in VVD, initiates the hinge to the N-cap (Figure 2.11). The N-caps of LOV2 and VVD are different, but

changes in both structures correlate with light-induced hydrogen bonding rearrangements within the hinge region [24]. In YtvA [23], the conserved Gln also flips upon adduct formation, but the peripheral hydrogen bonding network affected is again different from that of either phot1 LOV2 or VVD. In these systems, as with VVD, crystal packing constraints likely limit complete conversion to light-adapted conformations when dark-state crystals are irradiated. Nevertheless, we can conclude that core chemical mechanisms of the flavin and responses of immediately surrounding residues are conserved across the broad LOV family but the response of peripheral regions vary as sequences and structures diverge.

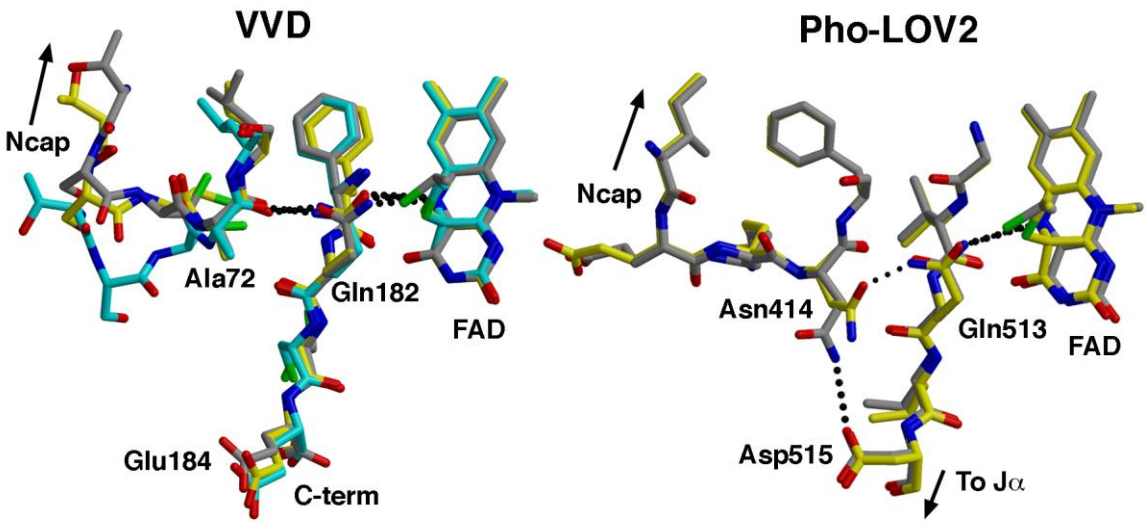


Figure 2.11: Comparison of the light induced changes in VVD and As LOV2. Comparison of the light induced changes in hydrogen bonding found for VVD (left) and the *Avena sativa* phototropin LOV2 domain (right). In both proteins, formation of the cysteinyl-flavin adduct causes similar changes to hydrogen bonding between the first and last □-strands of the respective PAS domains. In VVD, protonation of flavin N5 flips Gln182 in the light state (yellow and cyan for the two superimposed subunits) which results in a new hydrogen bond to the carbonyl of Ala72. In LOV2, the analogous Gln residue forms a similar hydrogen bond with the first residue of the first □-strand in the light state (yellow), but in this case, the acceptor is the side chain of Asn414, which no long hydrogen bonds to Asp515, as it does in the dark state

(grey). In the VVD LSD these changes result in restructuring of the linker loop leading to the Ncap. In the cyan subunit, the changes are larger than in the yellow subunit. No such large scale changes are seen in the LOV2 Ncap; however potential rearrangements are probably limited by the crystal lattice. *Avena sativa* LOV2 domains are shown for PDB codes 2V1A (dark state) and 2V1B (light state).

2.5 Methods

Protein expression and purification

Protein expression, purification and size exclusion chromatography of VVD-36II was carried out as previously described [22, 28].

Mutagenesis

The Tyr40Lys, Tyr40Glu, Tyr40Phe, Tyr40Ile, Thr69Trp and Ile52Cys variants were prepared by the QuickChange protocol (Stratagene) using the following primers. Tyr40Phe 5'- AGC CAT ATG CAT ACG CTC TTC GCT CCC GGC GGT-3', 5'- GTC ATA ACC GCC GGG AGC GAA GAG CGT ATG CAT-3'. Tyr40Lys 5'- AGC CAT ATG CAT ACG CTC AAG GCT CCC GGC GGT-3', 5'- GTC ATA ACC GCC GGG AGC CTT GAG CGT ATG CAT-3'. Tyr40Glu 5'- AGC AGC CAT ATG CAT ACG CTC GAG GCT CCC GGC GGT-3', 5'- GTC ATA ACC GCC GGG AGC CTC GAG CGT ATG CAT-3'. Tyr40Ile 5'- AGC CAT ATG CAT ACG CTC ATC GCT CCC GGC GGT-3', 5'- GTC ATA ACC GCC GGG AGC GAT GAG CGT ATG CAT-3'. Thr69Trp 5'-GAA CTG GGA CCT GTT GAC TGG TCA TGC GCT CTG-3', 5'-CAG AAT CAG AGC GCA TGA CCA GTC AAC AGG TCC-3'. Ile52Cys 5'- ATG GGC TAT CTG TGT CAG ATT ATG AAC AGG CCA-3', 5'- GTT CAT AAT CTG ACA CAG ATA GCC CAT AAT GTC-3', Thr69Cys 5'-GGA CCT GTT GAC TGC TCA TGC GCT CTG ATT CTG-3', 5'-CAG AGC GCA TGA

GCA GTC AAC AGG TCC CAG TTC-3'. Val86Cys 5'- GAC ACG CCA ATT TGC TAC GCC TCG GAA GCT TTT-3', 5'-TTC CGA GGC GTA GCA AAT TGG CGT GTC TTT TTG-3'. Ala101Cys 5'-GGA TAC AGC AAT TGC GAG GTC TTG GGG AGA AAC-3', 5'-CCC CAA GAC CTC GCA ATT GCT GTA TCC TGT CAT-3'. All the mutated genes were sequenced at the Biotechnology Resource Center of Cornell University.

Crystallization

VVD light-state crystals were grown overnight at 17 °C from VVD-36 II (*1*) under by vapor diffusion from a 5 µl drop containing 3 µl of protein at 3.5 mg/ml, including 5 mM DTT and 2 µl of reservoir solution. The protein was dissolved in a buffer containing 50 mM HEPES pH 8.0, 150 mM NaCl and 10% glycerol and the reservoir solution contained 11-15% PEG 6000, 0.1 M HEPES (pH=7.5 - 8.5), 5% MPD. The protein was exposed to white-light prior to crystallization and conversion to the light-adapted state was verified by UV-vis absorption spectroscopy. Crystal trays were exposed to white light once a day to maintain the protein in the light-adapted state. The largest crystals appeared as single plates (300 µm x 150 µm x 10 µm) that diffracted between 1.7 - 2.7 Å resolutions. Crystals were soaked in a cryoprotectant, consisting 20% v/v glycerol in reservoir solution, 30 s prior to flash cooling in liquid N_2. Monoclinic crystals of space group $P2_1$ were obtained, with two molecules per asymmetric unit.

Structure determination and refinement

Diffraction data was collected at 100 K with synchrotron radiation on the 24-ID-E beamline at the Advanced Photon Source (APS) at the Argonne National Laboratory, Chicago. The data was reduced and scaled with HKL2000 [35] (Table 2.1). Initial

phases for light-state VVD were obtained by molecular replacement by *Phaser* [36] using the dark-state structure (PDB ID: 2PD7) as a search model. Subsequent models were built in Xfit [37] and refined with CNS [38]. The flavin was left out of the molecular replacement probe and added at the later stages of refinement. The quality of the flavin electron density was used as a metric for improvement of the model. It was determined that the X-ray beam reduced the covalent adduct formed in light-adapted protein, therefore two strategies were employed to keep the adduct intact, variations of which have been applied previously to VVD [22] and other LOV proteins [39]. To reduce radiation exposure, each frame was collected over a 5° oscillation per second (X-ray flux of 2.33×10^{11} photons sec^{-1} per 20x20 mm^2 beam at 0.979 Å). Secondly, data were collected from ten different spots on a single crystal and then merged to generate the complete dataset.

VVD sequence analysis

Vivid sequence alignment and family analysis was performed using Pipealign (http://bips.u-strasbg.fr/PipeAlign/) [40]. Once a protein sequence is submitted, the software conducts a BLAST search and aligns the sequences, followed by refinement, validation and subfamily classification. In case of VVD, the resulting 170 sequences were classified under 10 groups, which contained various WC-1, GATA factors, Phototropins, Histidine kinases and putative blue-light receptors.

Construction of plasmids for phenotype analysis

Point mutations were introduced into the plasmid pCHC01-gfp-vvd-v5 as previously described [21] and confirmed by sequencing (Figure 2.1A).

Strains

The wild-type strain used here is OR74A. Knockout strains came from the *Neurospora* knockout project [41]. To establish that the *csr-1* knock-in strains used in this study were homokaryotic and carried only the transgene, we examined each strain by PCR analysis (30 cycles) to confirm the absence of the *csr-1* ORF in the genome [21, 42] (Figure 2.7B).

Culture conditions and light treatment

Culture procedures were performed as previously described [21]. After 24 h of culturing with constant shaking (125 rpm) in darkness (DD) at 25 ˚C, the flasks were moved to a shaker at 25 ˚C with continuous white light (LL), covering a wide-range of the spectrum from 400 nm to 700 nm (cool white fluorescent light bulb, GE F20T12-CW, 20 μmol/m^2/s). Mycelia were harvested before and after white light treatment.

Analysis of photoadaptation defects on slants

On day 0, conidia were inoculated onto a minimal slant and placed in an incubator with either constant white light (20 μmol/m^2/s) or constant darkness (DD) at 25 ˚C. Photographs were taken on day 4.

Analysis of carotenoid induction and RT-QPCR

These methods were performed as previously described [21].

REFERENCES

1. Taylor, B. and I. Zhulin, *PAS domains: internal sensors of oxygen, redox potential, and light.* Microbiology and Molecular Biology Reviews, 1999. **63**(2): p. 479.

2. Ayers, R. and K. Moffat, *Changes in Quaternary Structure in the Signaling Mechanisms of PAS Domains†,‡.* Biochemistry, 2008. **47**(46): p. 12078-12086.

3. Möglich, A., R.A. Ayers, and K. Moffat, *Structure and Signaling Mechanism of Per-ARNT-Sim Domains.* Structure, 2009. **17**(10): p. 1282-1294.

4. Crosson, S. and K. Moffat, *Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch.* The Plant Cell Online, 2002. **14**(5): p. 1067.

5. Christie, J., *Phototropin blue-light receptors.* Annual Review of Plant Biology, 2007. **58**: p. 21-45.

6. Zoltowski, B.D. and K.H. Gardner, *Tripping the Light Fantastic: Blue-Light Photoreceptors as Examples of Environmentally Modulated Protein-Protein Interactions.* Biochemistry, 2011. **50**(1): p. 4-16.

7. Heintzen, C., J.J. Loros, and J.C. Dunlap, *The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting.* Cell, 2001. **104**(3): p. 453-464.

8. Dunlap, J. and J. Loros, *The Neurospora circadian system.* Journal of biological rhythms, 2004. **19**(5): p. 414.

9. Heintzen, C. and Y. Liu, *The Neurospora crassa circadian clock.* Advances in Genetics, 2007. **58**: p. 25-66.

10. Krauss, U., et al., *Distribution and Phylogeny of Light-Oxygen-Voltage-Blue-Light-Signaling Proteins in the Three Kingdoms of Life.* Journal of bacteriology, 2009. **191**(23): p. 7234.

11. Dunlap, J. and J. Loros, *How fungi keep time: circadian system in Neurospora and other fungi.* Current opinion in microbiology, 2006. **9**(6): p. 579-587.

12. He, Q.Y., et al., *White collar-1, a DNA binding transcription factor and a light sensor.* Science, 2002. **297**(5582): p. 840-843.

13. Froehlich, A.C., et al., *White collar-1, a circadian blue light photoreceptor, binding to the frequency promoter.* Science, 2002. **297**(5582): p. 815-819.

14. Liu, Y., Q. He, and P. Cheng, *Photoreception inNeurospora: a tale of two White Collar proteins.* Cellular and Molecular Life Sciences, 2003. **60**(10): p. 2131-2138.

15. He, Q. and Y. Liu, *Molecular mechanism of light responses in Neurospora: from light-induced transcription to photoadaptation.* Genes & development, 2005. **19**(23): p. 2888.

16. Chen, C., et al., *Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora.* The EMBO Journal, 2009. **28**(8): p. 1029-1042.

17. Schwerdtfeger, C. and H. Linden, *Blue light adaptation and desensitization of light signal transduction in Neurospora crassa.* Molecular microbiology, 2001. **39**(4): p. 1080-1087.

18. Cheng, P., et al., *Functional conservation of light, oxygen, or voltage domains in light sensing.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(10): p. 5938.

19. Malzahn, E., et al., *Photoadaptation in Neurospora by Competitive Interaction of Activating and Inhibitory LOV Domains.* Cell, 2010. **142**(5): p. 762-772.

20. Hunt, S., et al., *VIVID interacts with the WHITE COLLAR complex and FREQUENCY-interacting RNA helicase to alter light and clock responses in Neurospora.* Proceedings of the National Academy of Sciences, 2010.

21. Chen, C.H., et al., *Physical interaction between VIVID and white collar complex regulates photoadaptation in Neurospora.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(38): p. 16715-16720.

22. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor vivid.* Science, 2007. **316**(5827): p. 1054-1057.

23. Möglich, A. and K. Moffat, *Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA.* Journal of Molecular Biology, 2007. **373**(1): p. 112-126.

24. Halavaty, A. and K. Moffat, *N-and C-Terminal Flanking Regions Modulate Light-Induced Signal Transduction in the LOV2 Domain of the Blue Light Sensor Phototropin 1 from Avena sativa†,‡.* Biochemistry, 2007. **46**(49): p. 14001-14009.

25. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch.* Science, 2003. **301**(5639): p. 1541-1544.

26. Harper, S.M., J.M. Christie, and K.H. Gardner, *Disruption of the LOV-J alpha helix interaction activates phototropin kinase activity.* Biochemistry, 2004. **43**(51): p. 16184-16192.

27. Purcell, E.B., et al., *An Analysis of the Solution Structure and Signaling Mechanism of LovK, a Sensor Histidine Kinase Integrating Light and Redox Signals.* Biochemistry, 2010. **49**(31): p. 6761-6770.

28. Zoltowski, B.D. and B.R. Crane, *Light activation of the LOV protein Vivid generates a rapidly exchanging dimer.* Biochemistry, 2008. **47**(27): p. 7012-7019.

29. Lamb, J.S., et al., *Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering.* Journal of the American Chemical Society, 2008. **130**(37): p. 12226-+.

30. Lamb, J.S., et al., *Illuminating Solution Responses of a LOV Domain Protein with Photocoupled Small-Angle X-Ray Scattering.* Journal of Molecular Biology, 2009. **393**(4): p. 909-919.

31. Zoltowski, B.D., B. Vaccaro, and B.R. Crane, *Mechanism-based tuning of a LOV domain photoreceptor.* Nature Chemical Biology, 2009. **5**(11): p. 827-834.

32. Chothia, C. and A.M. Lesk, *Helix Movements in Proteins.* Trends in Biochemical Sciences, 1985. **10**(3): p. 116-118.

33. Krissinel, E. and K. Henrick, *Inference of Macromolecular Assemblies from Crystalline State.* Journal of Molecular Biology, 2007. **372**(3): p. 774-797.

34. Cheng, P., et al., *WHITE COLLAR-1, a Multifunctional NeurosporaProtein Involved in the Circadian Feedback Loops, Light Sensing, and Transcription Repression of wc-2.* Journal of Biological Chemistry, 2003. **278**(6): p. 3801.

35. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode*, in *Macromolecular Crystallography, Pt A*. 1997. p. 307-326.

36. McCoy, A.J., et al., *Phaser crystallographic software.* Journal of Applied Crystallography, 2007. **40**: p. 658-674.

37. McRee, D.E., *XtalView Xfit - A versatile program for manipulating atomic coordinates and electron density.* Journal of Structural Biology, 1999. **125**(2-3): p. 156-165.

38. Brunger, A.T., *Version 1.2 of the Crystallography and NMR system.* Nature Protocols, 2007. **2**(11): p. 2728-2733.

39. Fedorov, R., et al., *Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii.* Biophysical Journal, 2003. **84**(4): p. 2474-2482.

40. Plewniak, F., et al., *PipeAlign: a new toolkit for protein family analysis.* Nucleic Acids Research, 2003. **31**(13): p. 3829-3832.

41. Colot, H.V., et al., *A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(27): p. 10352-10357.

42. Bardiya, N. and P.K. Shiu, *Cyclosporin A-resistance based gene placement system for Neurospora crassa.* Fungal Genetics and Biology, 2007. **44**: p. 307-341.

Chapter 3

MECHANISM OF LIGHT-SIGNALING IN *DROSOPHILA* CRYPTOCHROME

3.1 Abstract

Cryptochrome (CRY) is the principal photoreceptor in the *Drosophila melanogaster* circadian clock and plays non-photosensory roles in mammalian and other peripheral clocks. The C-terminus of *Drosophila melanogaster* cryptochrome (*Dm*CRY) is essential for signaling via light-induced interactions with downstream partners. Light reduces the flavin adenine dinucleotide (FAD) cofactor of the protein to and anionic semiquinone and leads to conformational changes at the C-terminal, but the mechanism is unclear. In the current study, we used limited proteolysis in combination with mass-spectroscopy and map the light-induced conformational changes to the C-terminus of CRY and its surrounding motifs. We show that light not only releases the C-terminal, but also rearranges the motifs around it to form the light-state conformation. Lifetime of the light-state conformation correlates well with the anionic semiquinone and chemical reduction of CRY in the dark also leads to the light-state conformation. The role of Trp536 residue and the C-terminal FFW motif were also studied. Based on the above data we propose a mechanism for light-induced conformational changes in *Drosophila* cryptochrome.

3.2 Introduction

Cryptochromes are flavoproteins found in all kingdoms of life [1-3]. They regulate growth and development in plants, act as photoreceptors in insect circadian clocks and are a necessary component of the mammalian clock, but their role in archea, bacteria and fungi is unknown [1, 4]. Cryptochromes have evolved from the photolyase family of proteins that repair UV-damaged DNA in a blue-light dependent fashion [3, 5]. Although some CRYs bind single stranded DNA, they no longer retain the ability to repair DNA. Based on their phylogenetic origin, CRYs can be broadly classified into three subgroups – plant CRY, animal CRY and CRY (DASH). Plant CRYs, like those of *Arabidopsis thaliana* CRY1, 2 (*At* CRY1, 2) are blue-light/near UV photoreceptor. Type I animal CRY are also photoreceptors, like *Drosophila melanogaster* CRY (*DmCRY)*, while *Homo sapien* CRY (*Hs*CRY) and Monarch butterfly CRY2 are type II CRY, and they primarily play a light-independent role as transcriptional repressors in the circadian clock [6]. The role of CRY (DASH) is unknown. Compared to photolyases, CRYs have varying C-terminal extensions that are essential for their function. All photolyases and CRYs have a flavin chromophore bound in an α/β domain and most also have a pteridine or deazaflavin cofactor that binds in a pocket closer to the N-terminus of the protein [5, 7, 8].

Found in all three kingdoms of life, photolyases are DNA repair proteins that utilize their blue-light sensitive FAD cofactor to repair DNA lesions. FAD is fully reduced to $FADH^-$ in the presence of blue light and energy transfer from the antenna cofactor, 5,10-methylenetetrahydrofolate (MTHF), via Förster resonance energy transfer (FRET) results in the excited $FADH^{-}$*. This excited $FADH^{-}$* is the catalytically active form that

repairs the DNA lesions by cyclic electron transfer [3, 9]. In plant CRYs, blue light reduces FAD to the biologically active neutral semiquinone, $FADH^{o}$, which re-oxidizes to FAD in the dark [1]. On the other hand, in *DmCRY*, blue-light reduces FAD to the anionic semiquinone (ASQ), FAD^{o-}, which oxidizes back to FAD in the dark. Some studies have suggested that the resting state of *DmCRY*, *in vivo*, is the ASQ and the observed oxidized form is an artifact of *in vitro* purification.

DmCRY is the principal photoreceptor in the Drosophila circadian clock, and it is involved in light-induced resetting of the clock [1, 10, 11]. The proteins Timeless (TIM) and Period (PER) inactivate the transcription factor complex CLOCK:CYCLE (CLK:CYC), which control the transcription of TIM and PER; this feedback loop forms the core clock. In the presence of blue light, CRY binds to TIM and sequesters it from the feedback loop by initiating its degradation [12-15]. CRY is also degraded in the presence of light and an E3 ubiquitin ligase, Jetlag (JET), is involved in both degradation processes [16]. The C-terminal extension of CRY is essential for its light-induced interactions with JET and TIM [17, 18]. Genetic and biochemical studies have shown that light induces conformational changes in CRY that lead to a higher affinity for its binding partners, and a CRY variant lacking 20-22 C-terminal residues is constitutively active [18-20]. There are currently two models for the mechanism of light-induced conformational changes in CRY. According to Model 1, FAD is in the oxidized state in dark and light reduces it to the ASQ, which leads to conformational changes and re-oxidizes back to FAD. According to Model 2, FAD is in the ASQ state in the dark and light produces a flavin excited state that then leads to the conformational changes in the C-terimus. Because CRY purifies in the oxidized state *in vitro*, according to Model

2, light has two roles: 1) to reduce FAD to the ASQ state and 2) excite the ASQ to the ASQ* state [20-22].

The structure of *DmCRY* reveals that the C-terminus is in close proximity to the catalytic FAD [8]. The C-terminal tail (CTT) (residues 528-539) occupies a groove analogous to the groove occupied by DNA substrate in photolyase, and it is perfectly poised to respond to the light-induced redox changes of the FAD. The CTT is connected to the photolyase homology domain (PHD), which ends at residue 516, via a linker corresponding to the residues 517-527. Motifs around the CTT, like the C-terminal lid (residues 420-446), the protrusion motif (288-306), the phosphate binding loop (249-263) and the CTT base loop (154-160), form the pocket for CTT and could be important for the execution of light responses (Figure 3.1). These motifs around the CTT along with the CTT will be referred to as the C-terminal tail coupled motif (CCM). The CTT is ~5 Å from the FAD and it could possibly respond to changes in the redox state of the FAD. The Phe534 of *Dm*CRY is in close proximity to the FAD and His378 lies in between the Trp536 and the FAD. Here, we use a limited proteolysis assay in combination with mass-spectroscopy, kinetic studies and mutational analysis to understand the mechanism of light-induced conformational changes in *DmCRY* and map them to the CCM.

3.3 Results

3.3.1 Light induces conformational changes in the CCM of CRY

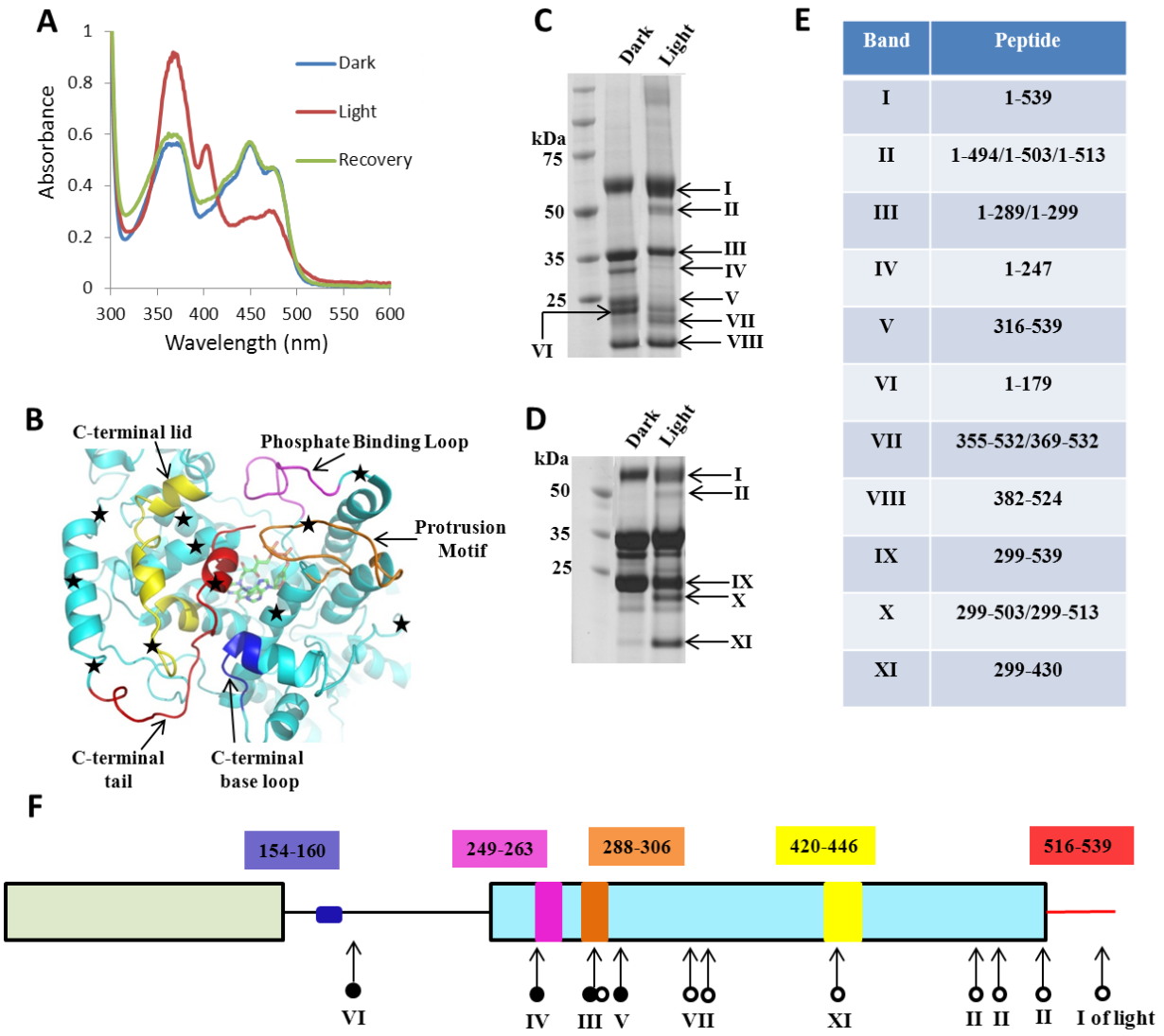
The C-terminal extension of *DmCRY* beyond the PHD is essential for its function as a photosensor. Correspondingly, the close proximity of the CTT to FAD (CTT of

DmCRY is ~5 Å from the FAD) provides the structural constraints necessary for this region of the protein to be modulated by flavin electronic states. In other blue-light sensors such as the light, oxygen and voltage (LOV) proteins and some CRYs, either the N-terminus or the C-terminus can be released in response to light. Thus, we hypothesized that the CTT could release (i.e. undock from the core protein) and adopt a new conformation in response to light. *Dm*CRY purifies with the flavin cofactor in the oxidized state, and light exposure reduces the FAD to an anionic semiquinone, which re-oxidizes in the dark (Figure 3.1). Ambient light does not affect the redox state of the protein; thus, a high intensity white light was used to generate the light state. All the dark state experiments were performed under a dim red light. To follow CTT release, we performed limited proteolysis with trypsin on dark and on light activated samples. A similar study was done on *Dm*CRY, where the authors used limited proteolysis to identify light-induced accessible sites and map them on the sequence of CRY [20]. They used Western Blots to analyze the samples and mapped the light-induced accessibility of the C-terminal extension. In light of the crystal structure of *Dm*CRY becoming available, we decided to revisit the assay and develop a simpler marker, using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) gels and not Western Blots, to follow the light-state conformation. We planned to use the simpler assay to test the kinetics of recovery and also the effect of various chemical reductants.

Trypsin digestion was performed for 30 s with 1:1 ratio of protein:trypsin concentration, before quenching, to obtain maximum light-state by avoiding competition between the rate of trypsin digestion and the reoxidation of CRY. The dark and light and recovery (re-oxidized in the dark for 1 h after light exposure) samples were digested

with trypsin and run on a SDS PAGE gel. Trypsin cleaves CRY differently in light compared to the dark, and the recovery sample looks more similar to the dark sample than the light exposed sample (Figure 3.1).

Mass spectral analysis of these differentially cleaved bands map the cleavage sites to in and around the CTT i.e. in the C-terminal tail coupled motif (CCM) (Figure 3.1). It should be noted that particularly for the higher molecular weight bands a mixture of closely related peptides can be found at each apparent molecular weight. Band I, present in all the samples, corresponds to the undigested protein (residues 1-539) while band II, present only in the light activated sample, corresponds to cleavage at Arg494. In the light sample, Band I was also shown to contain peptides corresponding to cleavage at Arg503, Arg513 and Arg530. Peptides with these C-termini are also found in Band II which only appears in the light sample and not in dark. Although peptides that differ by just a few residues do not resolve well within higher molecular weight bands, they are resolvable in the lower molecular weight bands: band IX is the peptide 299 - 539, while band X is a mixture of the peptides 299- 503 and 299-513. Residue Arg494 is at the N-terminus of the last helix of the PHD, while two other residues are within this helix (Arg503, Arg513), and one cleavage site is in the CTT itself (Arg530). This shows that the CTT and the last helix of the PHD are more accessible to proteolysis in light, indicative of a conformational change(s) leading to the release/undocking of the CTT. These sites are protected in the dark and in the recovery samples indicating a strong (or an absolute) dependence of the conformational change on the presence of the anionic semiquinone.



Band	Peptide
I	1-539
II	1-494/1-503/1-513
III	1-289/1-299
IV	1-247
V	316-539
VI	1-179
VII	355-532/369-532
VIII	382-524
IX	299-539
X	299-503/299-513
XI	299-430

Figure 3.1 Insights into the light-induced conformational changes of *Drosophila* Cryptochrome using mass-spectrometry. Limited proteolysis of the dark and the light exposed samples of Cry followed by mass-spectral analysis of the bands provides insight into the light induced conformational changes. (A) Electronic absorption spectra of the dark, the light and the recovered samples of Cry after light exposure. The FAD cofactor is in the oxidized state in dark, indicated by the absorption at 450 nm with two shoulders, and is reduced to an anionic semiquinone in light, indicated by peaks at 367 nm and 403 nm. The semiquinone is reoxidized back to the FAD in the recovery sample when kept in dark. (B) The crystal structure of Cry (PDB ID:3TVS) showing the C-terminal tail (CTT) in red, and the motifs around it; all these together are called the C-terminal motifs (CCM). The FAD is shown in green stick model behind the CTT. (C, D, E) SDS PAGE gels of the dark and light samples followed by limited proteolysis by trypsin. Panel C is a similar digestion at different protein concentrations.The bands indicated by arrows were analyzed by mass spectrometry and the corresponding peptide fragments are shown in panel E. Bands II, VII, X and XI are present only in the light, where as IV, V and VI are

present predominantly in the dark samples. This indicates a light induced reorientation leading to accessibility of certain sites and protection of others. The clevage sites are indicated by a 'star' in panel B. (F) The domain architecture of Cry along with the residues corresponding to the CCM are colour coded as shown in B. The cleavage sites are indicated by arrows and the open circles indicate accesibility to trypsin in the light, while the closed circles indicate accessibility in the dark. Ambiguity between light and dark is indicated by both the circles together. Light induced cleavage at Arg530 is nine residues shorter than the full length protein and could not be separated from Band I of the light sample. It is represented as 'I of light' with an open circle.

Band III is a mixture of peptides 1-289 and 1-299 and is present in all three samples, implying that Arg299 is quite accessible irrespective of light-induced conformational changes. Although the intensity of band III decreases in light compared to the dark, this effect is not consistent. Thus, Arg299, which is a part of the protrusion motif, could be slightly more accessible in the light, or there could be a site between residues 1 and 299 that is accessible in light, which leads to the degradation of the primary peptide responsible for band III in light. Band XI corresponds to a peptide of residues 299-430, which has two cleavage sites, Arg299 and Arg430, and is observed predominantly in light. As cleavage at Arg299 is light independent, it is the Arg430 in the C-terminal lid that must be accessible only in light. The lack of a peptide corresponding to 1-430 implies that Arg299 is much more accessible than Arg430.

Unlike the sites mentioned above, the cleavage sites corresponding to bands IV, V and VI are accessible in the dark and protected in the light. Band IV cleaves at Arg247, a residue in the phosphate binding loop, V is due to a cleavage at Arg316, in a helix connected to the protrusion motif and VI is due to a cleavage at Arg179, which is in the inter-domain linker. Light-induced conformational changes in *DmCRY* not only

release the CTT and render perspective cleavage sites accessible, but also rearrange the CCM, which protects certain sites and leads to the light-state conformation.

Band II will be used as a marker to follow the light-state conformation of CRY. To obtain better resolution of the band, gels were run for a longer time, which runs the lower molecular weight bands off the gels.

3.3.2 FFW motif in the C-terminal tail is important for CRY stability

The highly conserved residues Phe534, Phe535 and Trp536 in the CTT form the FFW motif (Figure 3.2). The FFW motif makes extensive contacts within the photolyase homology domain, which could be disrupted by redox changes in the FAD [8]. To study the role of this motif in signaling in cell culture, three variants were analyzed – Trp536Ala, Trp536Phe and one in which all three residues were mutated to alanine, referred to as the FFW variant *(These experimentswere performed by Deniz Top and Michael Young, The Rockefeller University)*. The stability of the FFW variant is severely affected in both the dark and in the light (Figure 3.2). A similar effect is seen on TIM stability with the FFW variant of CRY. Light-induced interaction between CRY and TIM leads to their degradation. Thus, mutating the FFW motif could affect its interactions with the core protein and predispose CRY to adopt a light state conformation even in dark, leading to its degradation. On the other hand, the Trp536Ala and Trp536Phe variants are more stable in the dark than the light and behave similarly to the wild type. Each variant is capable of stabilizing TIM in the dark state, much like WT, but TIM is more stable in the light with the two variants than with wild type. This indicates a

possible role for the Trp536 in TIM recognition/degradation, which is partially eliminated in the Trp536 variants.

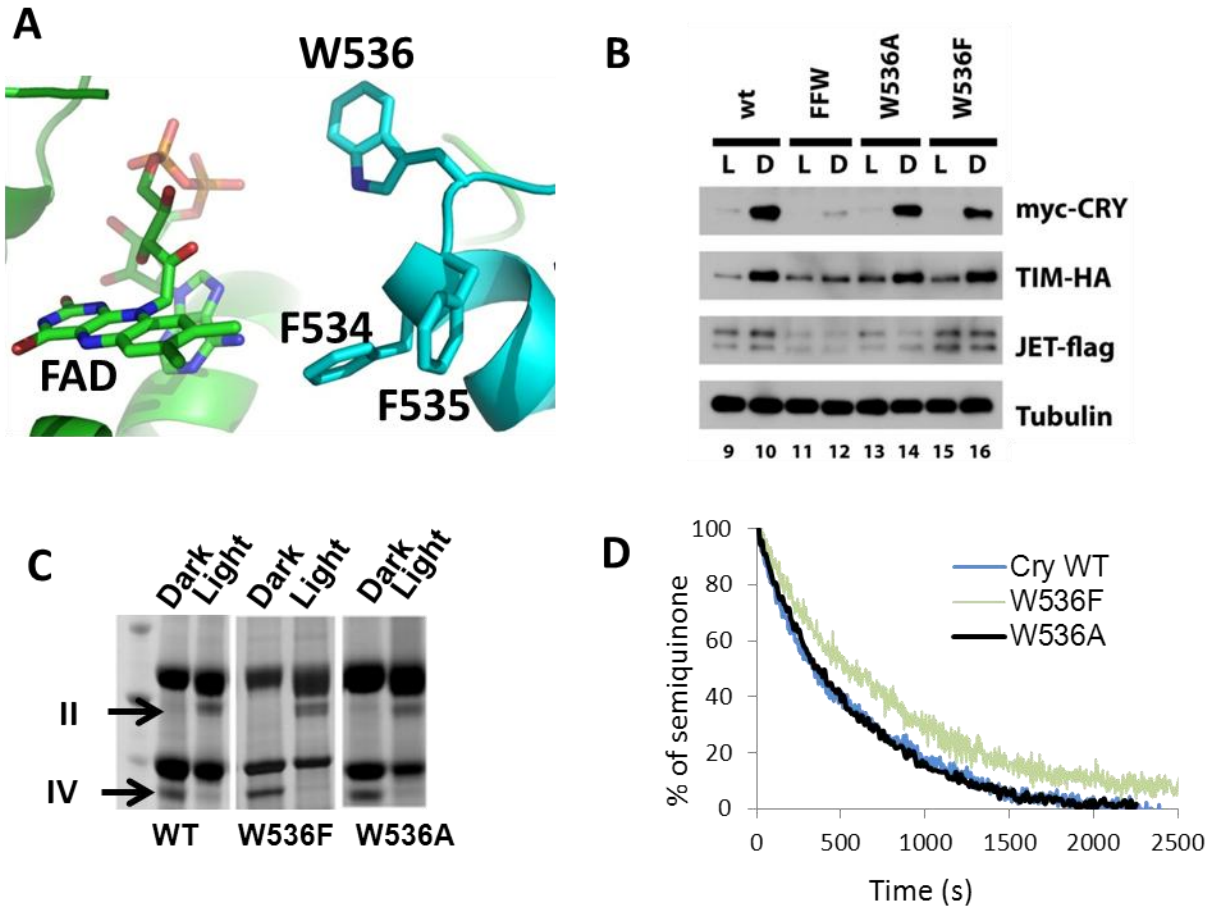


Figure 3.2 Role of W536 and the C-terminal FFW motif in light responses. (A) Crystal structure of *Dm*Cry shows close association of the F534, F535 and W536 (the FFW motif) residues in the C-terminal tail with the FAD (PDB ID:3TVS). (B) Western blot analysis of *Dm*Cry (Top panel) in S2 cells cotransfected with its binding partners, TIM (2nd panel) and JET (3rd panel) and the tubulin load control (bottom panel). L, cells were exposed to light for 1 hour before processing; D, cells were processed under red light. wt, wild-type; FFW, alanine mutation of amino acids 534-536. (C) SDS PAGE gels of the dark and light samples of wild type Cry and the W536 variants followed by limited proteolysis by trypsin. (D) The recovery kinetics of the anionic semiquinone back to the oxidized state in dark. The decay was followed at 403 nm, a peak which indicates the presence of the anionic semiquinone, for wild type (WT) Cry and the two W536 variants *(Experiments in panel B were performed by Deniz Top and Michael Young, The Rockefeller University).*

The Trp536Phe and Trp536Ala variants were also studied for light-induced conformational changes using the trypsin-digestion assay. Both of the variants behave similarly to the wild type CRY with the dark/light differences more pronounced in the Trp536Phe variant, which shows noticeably sharper bands (Figure 3.2). The rate of recovery of the anionic semiquinone formed under light back to the oxidized state was similar in the wild type and the two variants (Figure 3.2). This shows that the FAD redox activity is not affected by the two mutations. The *in vivo* and *in vitro* data show that the FFW motif affects the dark/light dynamics of CRY, while the Trp536Phe and Trp536Ala variants essentially behave like the wild type. Since the Trp536Phe provides sharper light/dark differences in the trypsin digestion assay, presumably due to a modestly weakened interaction of the CTT, it was used instead of wild type in several assays.

3.3.3 Rearrangement of the CCM depends on the lifetime of the ASQ

Using Bands II and IV as markers of the light-state conformation of CRY, we employed the trypsin digestion assay to probe the correlation between the light-induced reduction of FAD and the CRY conformational changes. To study the correlation between the semiquinone formation and the light-induced conformational changes, CRY was sequentially exposed to increasing light, and trypsin digestion assay was completed at the corresponding time points. Formation of the anionic semiquinone (ASQ), monitored at 403 nm, was gradually increased by varying the light exposure from 1 s to 120 s (Figure 3.3). ASQ concentration increases with increasing light up to 16s and remains constant thereafter. The intensity of band II grows (and band IV

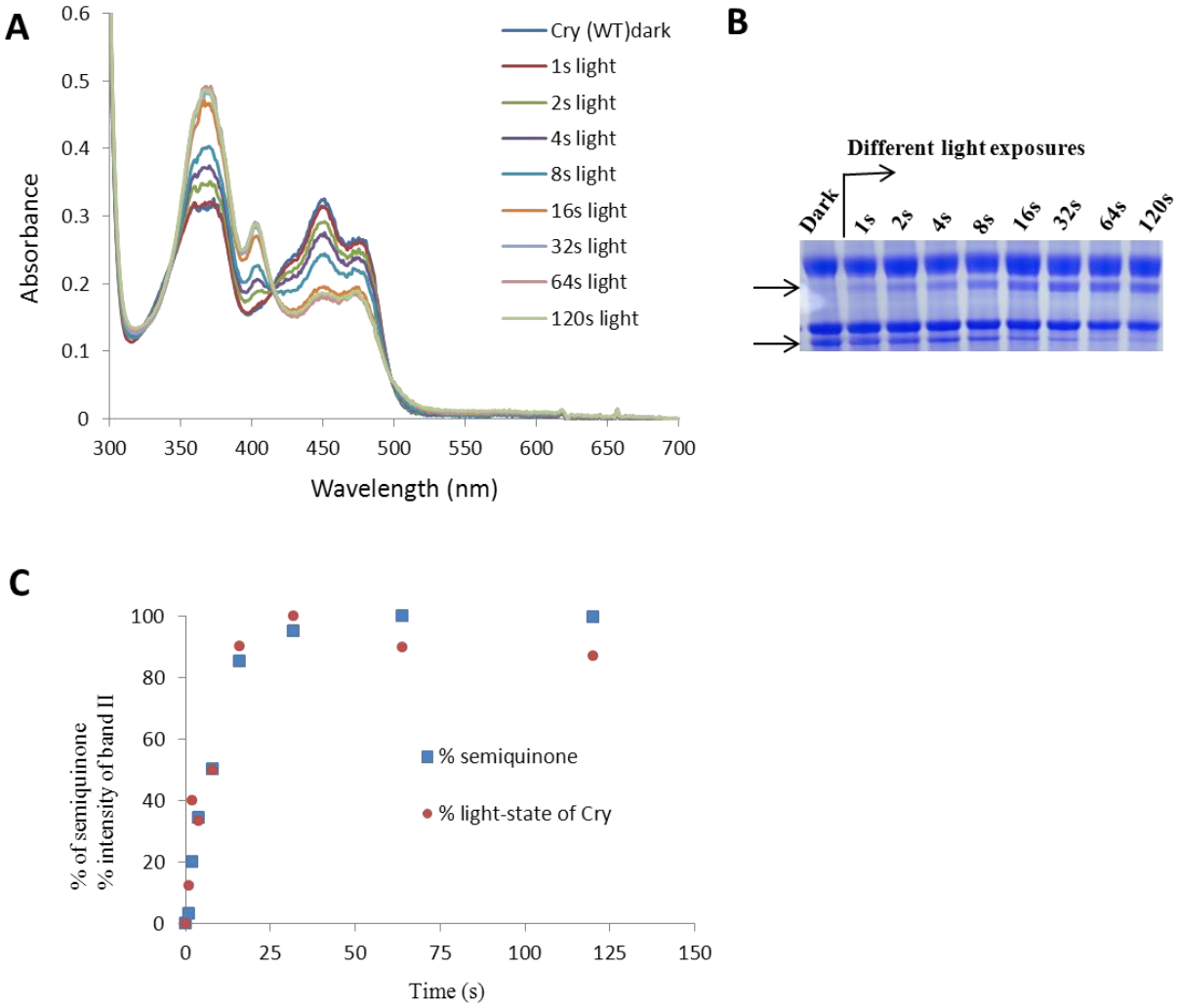


Figure 3.3 Correlation between ASQ formation and light-induced conformational changes (A) Absorption spectrum of *Dm*Cry with increasing light exposures showing an increase in the amount of the ASQ, indicated by the rise of 403 nm peak. (B) Trypsin digestion of *Dm*Cry with increasing light exposures, corresponding to those in panel A. Growth of band II and decay of band IV represent the conformational changes. (C) The rise in the absorbance at 403 nm and the growth in the intensity of band II represent the percentage of ASQ and the percentage of light-state conformation of Cry, respectively. Both proteolysis and spectral changes were followed until saturated and correlate well with each other.

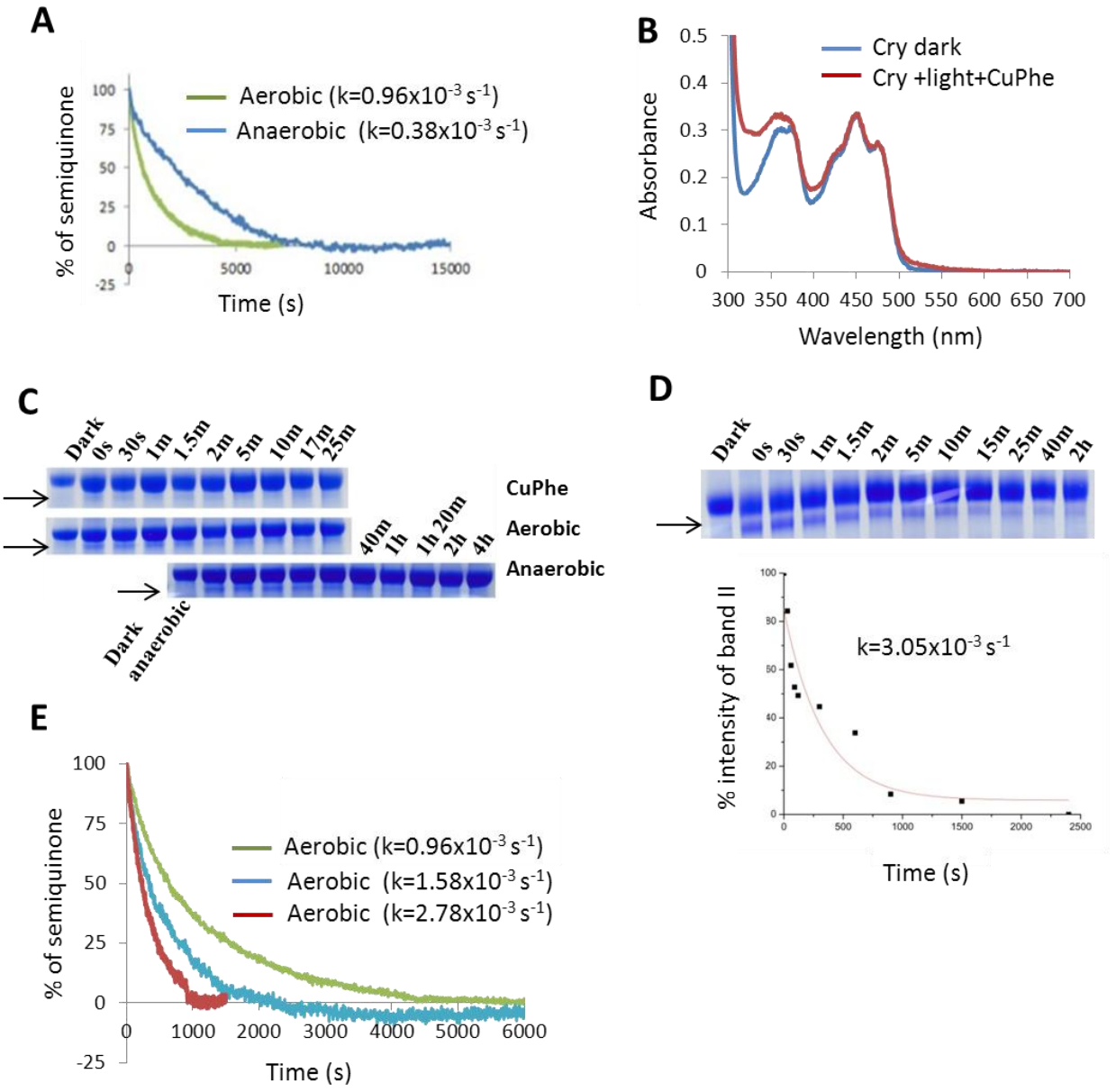


Figure 3.4 Correlation between the recovery of FAD and the light-state conformation (A) Recovery kinetics of the ASQ to the oxidized state of FAD in the presence and absence of oxygen. (B) Cu-Phenanthroline oxidizes the semiquinone within 1 m. (C) Recovery of the light-state conformation to the dark state is followed at three different rates of FAD recovery – Cu-Phenanthroline (fast), aerobic (medium) and anaerobic (slow). (D) The decay of the intensity of band II under aerobic conditions is plotted against time. (E) Aerobic recovery in different volumes i.e. at different concentrations of CRY (green curve is the maximum concentrations and red is the least).

decays) up to 16 s and then remains constant, which correlates strongly with the concentration of ASQ (Figure 3.3). To avoid the errors associated with such small exposure times and to avoid the ambiguity associated with the competition between trypsin digestion and the reoxidation of the FAD, we studied the kinetics of the recovery of CRY from light to dark (in combination with the trypsin digestion).

The recovery of *DmCRY* from the ASQ to the oxidized state occurs in 75 m with a $t_{1/2}$ of 12 m. This recovery is oxygen dependent and in anaerobic conditions, it occurs in 150 m with a $t_{1/2}$ of 30 m. Cu-Phenanthroline (Cu-Phe) oxidize the anionic semiquinone back to the dark state in 1 m (Figure 3.4). The rate of recovery of the CTT to the dark-state conformation was monitored in these three conditions by following the rate of loss of band II. Band II disappears faster in presence of Cu-Phe than in the aerobic condition and is preserved the longest in the anaerobic condition, which correlates well with the relative rates of recovery of the FAD in the three conditions (Figure 3.4). Thus, the formation of semiquinone, both in the forward, (upon exposure to light) and the reverse reactions (upon recovery) correlates completely with the CTT rearrangement. The rate constant for the aerobic recovery of the FAD is 0.003 s^{-1} and that of the CTT recovery is 0.001 s^{-1} are closely comparable indicating that the CTT reorientation depends on the lifetime of the semiquinone (Figure 3.4). Band IV also follows a similar trend in support of this conclusion.

To quantify the concentration dependence of the aerobic recovery, the kinetics was followed at different sample volumes; higher volumes of 1500 µl and 200 µl recovered faster than 100 µl (Figure 3.4). Thus, it is difficult to obtain a consistent value for the rate constant for aerobic recovery. Also, a direct comparison of the rate

constants from the proteolysis assay and the FAD recovery kinetics determined by absorption spectroscopy is not accurate.

3.3.4 Chemically reduced CRY is sufficient to adopt the signaling-state conformation

To distinguish between the two models of signaling in *DmCRY*, we completely eliminated light by chemically reducing CRY in dark and monitoring bands II and IV for 'light-state' or signaling-state conformation. Dithionite reduces CRY to an ASQ in the dark, the concentration of which peaks at 2 m after addition, and trypsin digestion at this time point reveals that the protein adopts a conformation that closely resembles the light-adapted state in terms of proteolytic sensitivity (Figure 3.5). Over the course of 2 h, dithionite completely reduces FAD to a two electron reduced state ($FADH^-$), which also adopts the signaling-state conformation albeit to a smaller extent, indicated by the lower intensity of band II at the 2 h time point (Figure 3.5). To check the reversibility of the reduction, after two hour dithionite reduction, the sample was kept in aerobic conditions (in the dark) for an hour followed by trypsin digestion. The protein recovers and adopts a dark state conformation. In some flavo-proteins dithionite chemically reacts with the FAD, and specific reducing agents could also affect the protein, so we tested another reducing agent, the Cr:EDTA complex (which as a potential of < -1.0 V vs. NHE) [23]. Addition of 11.5 mM Cr:EDTA reduces the FAD to an ASQ in 2 m and complete reduction takes 2 h. CRY rearranges to the signaling state in both conditions. Oxidation of Cr:EDTA reduced CRY causes behavior similar to dark state, and shining light on the reoxidized sample adopts the light-state (Figure 3.5). Thus, the data supports Model II in which the oxidized state of CRY is the ground state and the ASQ, formed either by light or chemically, is the signaling state which leads to conformational changes.

The concentration of the ASQ formed by both the reducing agents is less than that formed in the presence of light, as some of the CRY is completely reduced. The absorption peak of the reduced state of FAD is buried under the ASQ peak and hence cannot be observed until the ASQ concentration is low. The slower rate of formation of the $FADH^{-}$ could be due to a slow rate of FAD^{o-} protonation.

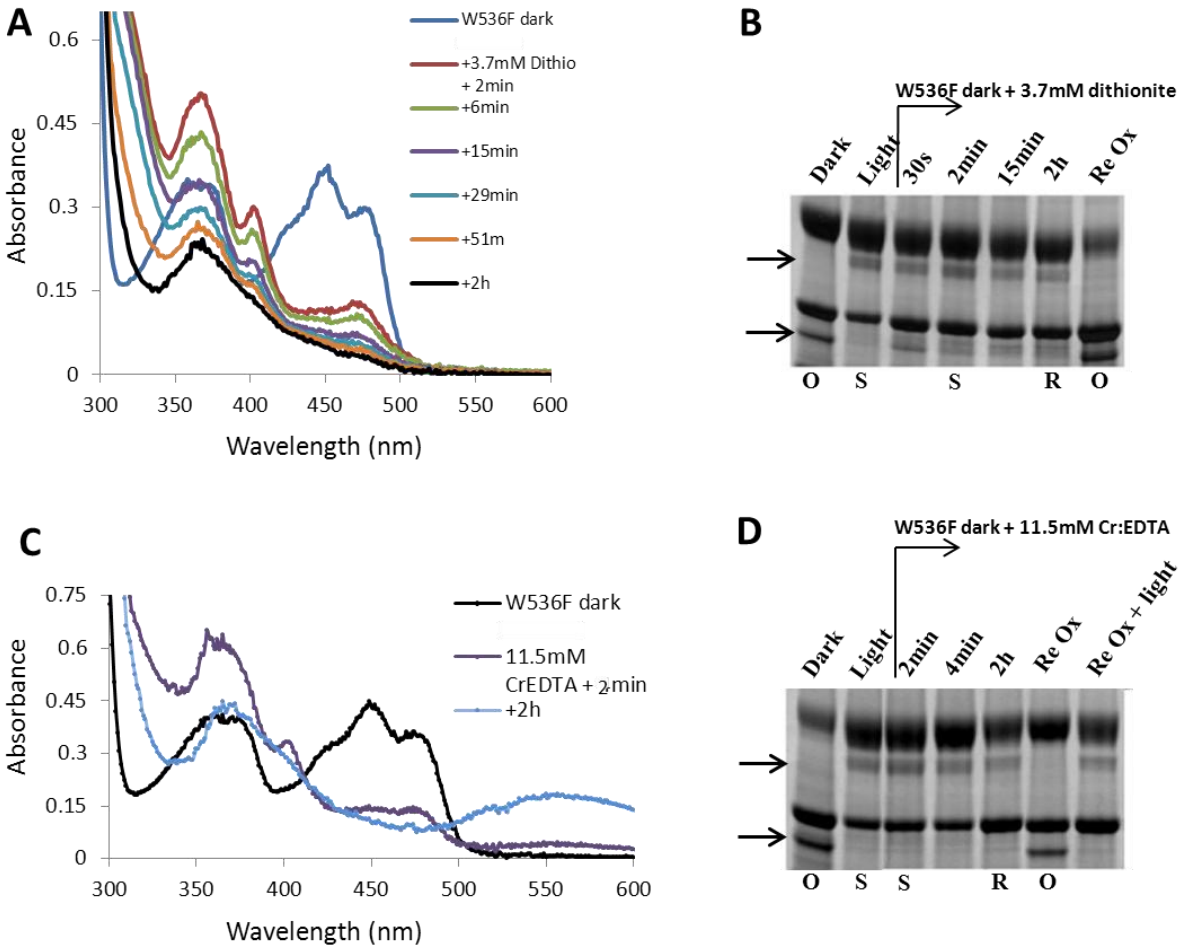


Figure 3.5 Chemical reduction of *Dm*Cry and the associated conformational changes (A, C) Reduction of W536F variant of CRY by dithionite (A) and Cr:EDTA (C), to anionic semiquinone (2 m after addition) and to completely reduced (2 hour after addition) states. (B, D) Trypsin-digestion at various time points after the addition of the reductants. ReOx, re-oxidation of the sample for 1 h after reduction by chemicals; ReOx + light, light exposure of the re-oxidized sample; Ox, oxidized; Sq, semiquinone; Rd, reduced. The lanes with Ox, Sq or Rd below them are completely in that particular redox state, while those that are not indicated are in a mixture of redox states.

3.4 Discussion

Drosophila cryptochrome is a well-studied member of the type I cryptochromes, but the mechanism of light sensing, the photocycle and the associated conformational changes are still a matter of debate [20-22]. *In vivo* electron paramagnetic resonance (EPR) studies of over expressed protein in insect cells suggest that the oxidized state of Cry is the ground state while the ASQ is the light excited state, supporting Model 1 [24]. Mutating residues in the conserved Trp-triad, which provides the electron for reduction, blocks the reduction *in vitro* but not *in vivo [20, 25, 26]*. This could indicate that the ASQ is the ground state of CRY in cells, eliminating the need for electrons, which supports the Model 2. Recently, it was shown that dithionite reduces *DmCRY* to an ASQ in the dark and this does not lead to conformational changes, but dithionite reduction followed by light exposure leads to the expected conformational changes [20]. These results add weight to Model 2. We performed a similar experiment in the dark, to probe the role of light and show that dithionite and Cr:EDTA not only reduces CRY to an ASQ, but also to a two electron reduced state. Surprisingly, both the 1-electron and the 2-electron reduced states restructure the C-terminal motif similar to light. The chemically reduced samples were re-oxidized and the trypsin digestion pattern was similar to dark sample. Light exposure of the re-oxidized sample adopted a light-state conformation, which indicated no damage to the protein by the reducing agents. This clearly supports Model 1 of light sensing in *DmCRY*. At the least we can conclude that formation of the anionic semiquinone causes conformational changes in CRY that increase the exposure of the CTT and rearrange structural elements in close proximity of its bound state. The recovery kinetics data also adds weight to Model 1; the lifetime of the ASQ determines

the lifetime of the light-state conformation. The excited ASQ of Model 2 would be too short-lived to affect the recovery of the light-induced conformational changes.

Our mass-spectral analysis agrees well with previously observed CRY cleavage sites (503 and 420) that are accessible in the light. We identified more light-accessible cleavages sites, in the C-terminal tail (530) and the helix (494) connected to it, which indicate a light induced release/reorientation of ~45 C-terminal residues. We also identified sites that are accessible in dark and protected in the light and mapped them to the C-terminal motif. These results suggest that light leads to a complex reorientation of the C-terminal motif and not just a release of the C-terminal tail. This light-induced reorientation could create a new interface for recognition of TIM and JET [18].

The limited proteolysis assay is a good marker to assay the light state conformation of Cry . Using this assay, we determined that the Trp536 residue, alone, is not critical for adopting the light-state conformation. *In vivo* studies confirm this conclusion and show that the entire FFW motif is critical for the functioning of Cry. The *in vivo* data also hints at a more active role for the C-terminal tail, specifically the Trp536, as a recognition motif for TIM.

3.5 Methods

Growth and purification of *Dm*CRY

*Dm*CRY was overexpressed in insect cells and purified under red light as described previously [8].

UV-visible absorption spectrum

A 0.2 cm path length quartz cuvette with a sample volume of 13 ul was used for all the spectral analysis. The dark spectrum was obtained by adding the sample to the cuvette in the dark and covering it with aluminium foil while recording the spectrum. The light from the spectrophotometer is the only light the sample is exposed to and it is not sufficient to alter the redox state of the FAD. The light-spectrum was obtained by shining white light (Oriel 600 lamp) for 2 m and taking the spectrum, while the re-oxidized spectrum was obtained by letting the protein recover completely in dark for 2 h and recording the spectrum using an aluminium foil as mentioned above for the dark spectrum. All the spectra were normalized by subtracting the absorption at 700 nm. The spectra in Fig 3.3A were obtained by shining light with increasing exposure time and taking a spectrum after each exposure; after the dark spectrum, the sample was exposed to 1 s of light followed by a spectrum, then 1 s of light (total of 2 s) followed by a spectrum, 2 s of light (total of 4 s) followed by a spectrum and so on up to a total of 120 s of light exposure. The oxidation spectrum of Cu-Phe, in Fig 3.4B, was obtained by exposing 11 ul of CRY to 2 m of light and adding 1 ul of 1.2 mM Cu-Phe to it.

Recovery kinetics of CRY

A 1.0 cm path length quartz cuvette with a sample volume of 80 ul was used for the kinetic analysis. The samples was exposed to white light for 2 m and the absorption at 367 nm and 403 nm were followed to monitor the decay of the anionic semiquinone, while the absorption at 450 nm and 475 nm were followed to monitor the growth of the oxidized state. All these were normalized by subtracting the absorption at 700 nm. The

anaerobic kinetics was obtained by preparing the samples under anaerobic conditions in the glove bag. The cuvette with CRY was passed through 12 cycles of vacuum and nitrogen to remove the oxygen and taken into the glove bag. As the spectrophotometer was outside the glove-bag in aerobic conditions, the cuvette was sealed using parafilm to maintain the anaerobic condition in the cuvette. The sample was exposed to 2 m of light and the kinetics was performed as mentioned above.

Preparation of the Cr:EDTA complex

$Cr(II)Cl_2$ is shipped under anaerobic (under nitrogen) conditions. This was taken in to the glove bag and 5-20 mg was weight out for every reaction (in the glove bag). 0.5 M HCl and 0.5 M EDTA solution (pH=8.0) were degassed to remove dissolved oxygen and taken in to the glove bag. A 0.5 M solution of $Cr(II)Cl_2$ was prepared using 0.5 M HCl, which is cyan in colour. To 15 ul of 0.5 M $Cr(II)Cl_2$, 35 ul of 0.5 M EDTA solution was added and the colour stays very close to cyan. This forms a 150 mM Cr:EDTA complex. Over the course of 3 h, the colour changes from cyan to pale royal blue to darker royal blue to pale purple (Figure 3.6). The pale purple is the active form which was used in this study and the colour stays for a little more than an hour. It slowly becomes darker purple, which is also active (but lesser than the pale purple), but a day later it becomes dark blue which is useless.

CRY reduction with dithionite and Cr:EDTA

Dithionite war weighed in the glove bag and a stock was prepared in degassed buffer. The buffer is same as the one used for the protein. A 0.2 cm path length

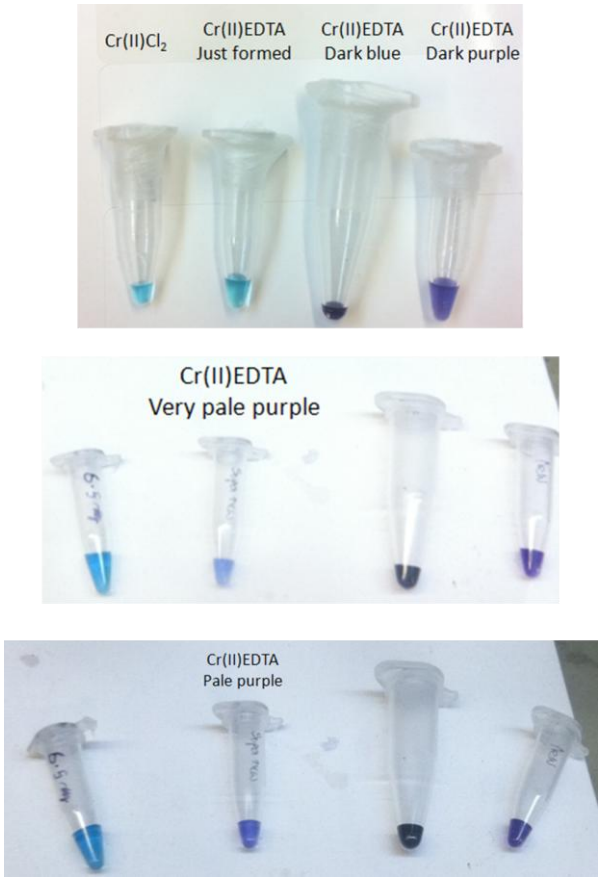


Figure 3.6 Colour changes associated with Cr:EDTA. The Cr:EDTA complex changes from cyan to pale purple over the course of 3 h and the pale purple complex is the active form. The dark purple on the far right is less active and the dark blue is inactive as a reducing agent.

cuvette with 13 ul CRY was used to take a dark spectrum and was then passed through 12 cycles of vacuum and nitrogen to remove the oxygen before taking into the glove bag. To 13 ul of CRY, 1 ul of 52 mM dithionite was added under anaerobic conditions and as the spectrophotometer was outside the glove-bag in aerobic conditions, the cuvette was sealed using parafilm to maintain the anaerobic condition in the cuvette. Successive spectra were taken at different time points.

For Cr:EDTA reduction, the protocol was same as the dithionite reduction. To 12 ul of CRY, 1 ul of 150 mM Cr:EDTA was added under anaerobic conditions in the glove-bag and sealed with parafilm to maintain anaerobic condition inside the cuvette.

Trypsin digestion assay

For all the samples that were analyzed using the limited proteolysis by trypsin, the following protocol was employed: to 2 ul of 200 uM CRY, 2 ul of 200 uM trypsin was added and the reaction was quenched after 30 s with 2 ul of 400 uM trypsin inhibitor and 10 ul of SDS followed by 10 min heating at 90 ^{o}C. 5 ul of the above solution was loaded on the gel. Then concentration of CRY, trypsin and trypsin inhibitor were kept constant in the entire study and unless mentioned, they are as above. For the light-samples, CRY was exposed to 2 m of white light (Oriel 600 lamp) before adding trypsin and the 30 s digestion was also performed under the same light. For the dark-samples, the reaction was carried out under red light. The intensities of the bands were quantified using ImageJ.

The forward kinetics of the light-state conformation in Figure 3.3B was followed by aliquoting 2 ul of CRY in eight tubes and exposing them to different lengths of time from 1 s to 120 s, before starting the proteolysis. During the 30 s of digestion, the samples were kept under red light.

The aerobic recovery kinetics of the light-state conformation in Fig 3.4C and 3.4D was followed by taking two aliquots of 10 ul CRY in each tube. Tube-1 was exposed to 2 m of white light and 2ul aliquots was taken out at various time points (30 s, 1.5 m, 5 m and 17 m) for the reaction. At the mentioned time point, 2 ul of trypsin was added and

the above mentioned protocol was followed. Tube-2 was then exposed to 2 m of white light and 2ul aliquots was taken out at various time points (1 m, 2 m, 10 m and 25 m) for the reaction just as in tube-1. The anaerobic samples in, Figure 3.4C, were prepared similar to the aerobic ones but under anaerobic condition in the glove-bag. Tube-1 was used to get the 2 m, 10 m, 25 m, 1 h and 2 h time points, while tube-2 was used to get 5 m, 17 m, 40 m, 1 h 20 m and 4 h time points. In both the aerobic and the anaerobic conditions, the dark and the 0 s time points were obtained using separate samples similar to other dark and light samples. The kinetics under Cu-Phe condition, in Fig 3.4C, was performed at similar time points as the aerobic condition. 23 ul of CRY was exposed to 2 m of light and 1 ul of 2.4 mM Cu-Phe was added. The total 24 ul was divided into two tubes and the protocol used for aerobic analysis, mentioned above, was repeated.

For the trypsin digestion after chemically reducing CRY in Fig 3.5B & 3.5C, the reducing agent was added to CRY and aliquots were taken at various time points for trypsin digestion. The entire experiment was performed under red light. To 13 ul of CRY, 1 ul of 52 mM dithionite was added under anaerobic conditions in the glove-bag and 2 ul aliquots were taken a 30 s, 2 m, 15 m and 2 h. These samples were treated with trypsin for 30s and quenched as mentioned above. The ‘Re Ox’ sample was obtained by taking an aliquiot at 2h out of the glove bag and exposing to aerobic conditions for 1 h, to re-oxidize the FAD to the dark state. The sample was then digested with trypsin. The dark and light samples were obtained by diluting 13 ul of CRY with 1 ul buffer and performing the trypsin digestion as mentioned earlier. For the Cr:EDTA samples, to 12 ul of CRY, 1 ul of 150 mM Cr:EDTA was added under

anaerobic conditions in the glove-bag and 2 ul aliquots were taken a 2 m, 4 m and 2 h. These samples were treated with trypsin for 30s and quenched as mentioned above. Two aliquiots were taken out of the glove bag at the 2 h time point and exposed to aerobic conditions for 1 h, to re-oxidize the FAD to the dark state. One of the samples was then digested with trypsin to get the ‘Re Ox’ time point, while the other was exposed to 2 m light before digestion to get the ‘Re Ox + light’ time point. The dark and light samples were obtained by diluting 12 ul of CRY with 1 ul buffer and performing the trypsin digestion as mentioned earlier.

REFERENCES

1. Chaves, I., et al., *The cryptochromes: blue light photoreceptors in plants and animals.* Annual review of plant biology, 2011. **62**: p. 335-364.
2. Peschel, N. and C. Helfrich-Förster, *Setting the clock-by nature: circadian rhythm in the fruitfly Drosophila melanogaster.* FEBS letters, 2011.
3. Sancar, A., *Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors.* Chemical Reviews-Columbus, 2003. **103**(6): p. 2203-2238.
4. Mohawk, J.A., C.B. Green, and J.S. Takahashi, *Central and Peripheral Circadian Clocks in Mammals.* Annual Review of Neuroscience, 2012(0).
5. Sancar, A., *Photolyase and cryptochrome blue-light photoreceptors.* Advances in protein chemistry, 2004. **69**: p. 73-100.
6. Tomioka, K. and A. Matsumoto, *A comparative view of insect circadian clock systems.* Cellular and molecular life sciences, 2010. **67**(9): p. 1397-1406.
7. Selby, C.P. and A. Sancar, *The Second Chromophore in Drosophila Photolyase/Cryptochrome Family Photoreceptors.* Biochemistry, 2011.
8. Zoltowski, B.D., et al., *Structure of full-length Drosophila cryptochrome.* Nature, 2011. **480**(7377): p. 396-399.
9. Li, J., et al., *Dynamics and mechanism of repair of ultraviolet-induced (6-4) photoproduct by photolyase.* Nature, 2010. **466**(7308): p. 887-890.
10. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-679.
11. Emery, P., et al., *Drosophila cryptochromes: a unique circadian-rhythm photoreceptor.* Nature, 2000. **404**(6777): p. 456-457.
12. Hardin, P.E., *5 Molecular Genetic Analysis of Circadian Timekeeping in Drosophila.* Advances in Genetics, 2011. **74**: p. 141.
13. Glossop, N.R.J., L.C. Lyons, and P.E. Hardin, *Interlocked feedback loops within the Drosophila circadian oscillator.* Science, 1999. **286**(5440): p. 766.
14. Saez, L., P. Meyer, and M. Young. *A PER/TIM/DBT interval timer for Drosophila's circadian clock*. 2007. Cold Spring Harbor, NY.

15. Meyer, P., L. Saez, and M.W. Young, *PER-TIM interactions in living Drosophila cells: an interval timer for the circadian clock.* Science's STKE, 2006. **311**(5758): p. 226.
16. Koh, K., X. Zheng, and A. Sehgal, *JETLAG resets the Drosophila circadian clock by promoting light-induced degradation of TIMELESS.* Science's STKE, 2006. **312**(5781): p. 1809.
17. Busza, A., et al., *Roles of the two Drosophila CRYPTOCHROME structural domains in circadian photoreception.* Science's STKE, 2004. **304**(5676): p. 1503.
18. Peschel, N., et al., *Light-dependent interactions between the Drosophila circadian clock factors cryptochrome, jetlag, and timeless.* Current biology, 2009. **19**(3): p. 241-247.
19. Dissel, S., et al., *A constitutively active cryptochrome in Drosophila melanogaster.* Nature neuroscience, 2004. **7**(8): p. 834-840.
20. Ozturk, N., et al., *Reaction mechanism of Drosophila cryptochrome.* Proceedings of the National Academy of Sciences, 2011. **108**(2): p. 516-521.
21. Liu, B., et al., *Searching for a photocycle of the cryptochrome photoreceptors.* Current opinion in plant biology, 2010. **13**(5): p. 578-586.
22. Partch, C.L. and A. Sancar, *Photochemistry and Photobiology of Cryptochrome Blue-light Photopigments: The Search for a Photocycle.* Photochemistry and photobiology, 2005. **81**(6): p. 1291-1304.
23. Crane, B.R., L.M. Siegel, and E.D. Getzoff, *Structures of the siroheme-and Fe4S4-containing active center of sulfite reductase in different states of oxidation: heme activation via reduction-gated exogenous ligand exchange.* Biochemistry, 1997. **36**(40): p. 12101-12119.
24. Hoang, N., et al., *Human and Drosophila cryptochromes are light activated by flavin photoreduction in living cells.* PLoS biology, 2008. **6**(7): p. e160.
25. Öztürk, N., et al., *Animal Type 1 Cryptochromes.* Journal of Biological Chemistry, 2008. **283**(6): p. 3256-3263.
26. Song, S.H., et al., *Formation and function of flavin anion radical in cryptochrome 1 blue-light photoreceptor of monarch butterfly.* Journal of Biological Chemistry, 2007. **282**(24): p. 17608.

Chapter 4

PHOTORECEPTION IN CRY AND LOV PROTEINS AND ITS IMPLICATIONS

4.1 Conclusions

Light-induced conformational changes in photoreceptors leads to new protein:protein interactions essential for signal transduction [1]. In this thesis we studied the mechanism of light-induced conformational changes in a fungal LOV protein and in fruit fly cryptochrome. We also studied the mechanism of dimerization in the LOV protein. Earlier studies on understanding the light-induced structural changes in phototropin LOV domains, YtvA LOV domain and VVD were achieved by exposing the dark-grown crystals to light and collecting X-ray diffraction data from these crystals [2-4]. Modest changes in the vicinity of the flavin and in some loops were observed, but larger conformational changes are constrained by crystal packing. Nuclear magnetic resonance (NMR) studies on isolated phototropin LOV domains indicate a light induced undocking of a C-terminal helix from the LOV core leading to activation of the attached kinase domain [5]. The question of how adduct formation leads to conformational changes still remained unanswered. In chapter 2, we crystalized the fully light-adapted state of VVD and determined the structure to 2.7 Å resolution [6]. Comparing this with the structure of VVD in the dark-state, we observe a large conformational change in the N-terminal tail that undocks from the protein core. The pathway within the protein from adduct formation to conformational change at the N-terminus involves new hydrogen bonding networks. The protein crystalizes as a dimer with the newly undocked N-terminal tail in the dimer interface. Mutational analysis of residues in the N-terminal tail

confirms the crystallographic dimer as the solution state dimer and testing these variants in *Neurospora* confirmed this as the signaling state dimer. The structure fits well with all the previous studies on the dimerization of Vivid. This is the first fully light-adapted LOV structure of a biologically relevant oligomer.

Cryptochrome is the principal photoreceptor in *Drosophila melanogaster* [7-9]. We recently determined the structure of *Dm*CRY which is the first structure of an animal CRY and it provides clues on the mechanism of light signaling [10]. In chapter 3, we developed a limited proteolysis assay, using trypsin, to follow the light state conformation of *Dm*CRY. A few cleavage sites in the protein are accessible in light, while a few others are protected. Using extensive mass-spectral analysis, we mapped the cleavage sites to the C-terminal tail and the surrounding motifs. It was known that the C-terminus is important for the functioning of *Dm*CRY and it would undock from the protein core in response to light. Our data not only shows that the C-terminal tail is released in light, but also a complex rearrangement of the surrounding motifs lead to the light-state conformation of the protein. Studying the kinetics of formation and recovery of the light-state conformation, using the proteolysis assay, we show a strong correlation between the redox state of the FAD and the conformational changes. Using chemical reductants to reduce CRY in dark, we confirm that the role of light is just to reduce the FAD and that the light-induced conformational changes of *Dm*CRY depend only on the redox state of the FAD. The role of light in *Dm*CRY signaling is still under debate and we believe that our data strongly argues that the role of light is to reduce the flavin cofactor.

4.2 Discussion

Light entrainment is essential for organisms to reset their circadian clocks to the 'correct time'. This helps them to get the most out of sunlight and receive environmental cues essential for survival. Sensing the blue light seems to be critical for many organisms, which could be due to a combination of factors – 1) higher penetration of the blue wavelength in water/oceans 2) the higher energy in the blue photons and 3) its wavelength being closer to UV rays; it could be an indicator for the presence of the damaging UV radiations. Nature has two small molecules, flavin and p-coumaric acid, distributed in four protein domains to sense blue-light. Photoactive yellow proteins (PYP) use the p-coumaric acid while the light, oxygen or voltage sensing (LOV) domain, the cryptochrome/photolyase (CRY/PL) proteins and the sensory proteins of blue-light using flavin (BLUF) use flavin. In the current work we studied LOV and CRY proteins to understand their mechanism of light sensing and signaling in the context of circadian rhythms.

The LOV domain is found in bacteria, archaea, prostita, plants and fungi, but not in animals, whereas CRY is found in all kingdoms of life [1, 11-13]. Lower organisms like bacteria and fungi depend extensively (maybe exclusively) on the LOV domain to sense light while animals depend on CRY for the photo input to the circadian clock. Mammals most likely use the G-protein receptor melanopsin for circadian light input [14]. Plants use both LOV and CRY to sense light and do not seem to depend on any one principal photoreceptor [15]. Plants depend on sunlight for their survival and it is advantageous to spread out light-sensing over various proteins, rather than depend on only one photoreceptor.

It is unclear why animals do not have a LOV domain and have a CRY instead. The higher complexity of the animals and the development of various peripheral clocks could be the possible reasons that tilted the balance in favour of CRY. Though the core clock in most organisms is based on a transcriptional-translational feedback loop, the animal clocks have more components, i.e. more inputs, more outputs and more feedback loops, than the fungal clock. With an increasing complexity of organisms, there arose a necessity to sense more signals than just light. The development of peripheral clocks, in various parts/organs of organisms, needed sensors for inputs other than light. Having one versatile sensor is more efficient than one for each function. In the animal kingdom, CRY acts as a photosensor, a magnetosensor, an energy sensor and a core clock protein (refer to Chapter 1 for more details).

Why did nature tweak CRY/PL, and not the LOV domain, to get a versatile sensor? The orientation of the flavin could explain this (Fig. 4.1). In LOV domain, there are two types of flavins cofactors – 1) the flavin mononucleotide (FMN) and 2) flavin adenine dinucleotide (FAD). The adenine moiety is ~15 Å away from the isoalloxazine ring in the FAD, while the FMN does not have an adenine. In LOV proteins, the flavin shuttles between the oxidized and the two-electron reduced states in response to dark/light changes. The one-electron reduced state is transiently formed and seems to have no signaling role. On the other hand, CRYs always have a FAD cofactor in a 'U-shaped' conformation with the adenine ~6 Å from the isoaloxazine ring, and both the one-electron and the two-electron reduced state of the FAD in CRY have signaling roles. The adenine moiety, which is in the close proximity of the isoalloxazine ring, could stabilize the one-electron reduced states. As the flavin in LOV domain shuttles between

just two redox states, LOV acts like an ‘On/Off’ switch, while the CRY can be finely modulated over four different redox states of the FAD and acts as more than just a switch. Owing to this fine modulation, CRY could act as a versatile sensor.

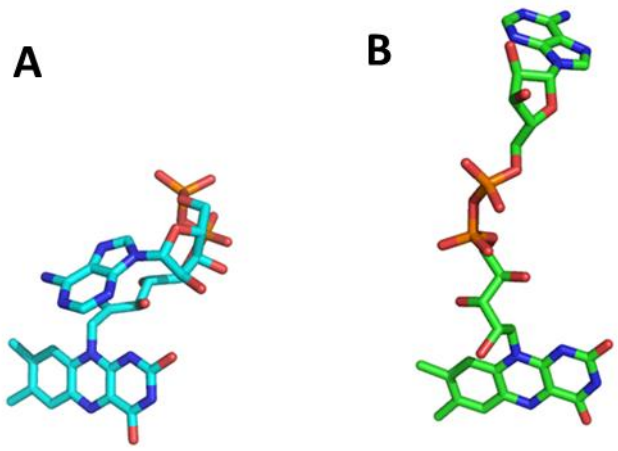


Figure 4.1 Comparison of FAD in the CRY/PL family and the LOV family. (A) The ‘U-shaped’ conformation of FAD in CRY/PL and (B) the FAD conformation in LOV proteins

The role of a clock is different in lower organisms, like *Neurospora*, than in higher organisms, like fruit fly or mammals. In *Neurospora*, the clock could help time gene duplication, for cell division, during the night. Hostile environmental conditions involving UV rays during the day could damage the DNA and sensing sunlight was critical for survival (moisture availability could be another factor influencing sporulation). Hence, an ‘on/off’ switch in the form of a LOV domain was necessary and sufficient. Higher organisms evolved under the ozone layer and sensing the sun was not critical for survival. The role of the clock is as much to sense and synchronize various activities

within the body as to sense the day/night cycle. Hence, they require not just an ‘On/Off’ switch for light but also need to sense various other cues and CRY fits the bill.

As the complexity of the clock increases, it requires more protein:protein interactions. LOV domains are usually a part of multi-domain proteins with its effector domain/protein being a part of the same protein. They seem to dimerize only with other LOV domains and hence are limited in their ‘reach’. On the other hand, CRY could interact with a large number of proteins – TIMELESS, JETLAG, PERIODs, CRYs, CLOCK, CYCLE, BMAL 1, AMPK and other kinases, allowing for a more complex and versatile network.

4.3 Future Directions

<u>To understand the role of VVD Tyr40 mutations in WC-1</u>

The Tyr40 residue in the N-terminal tail of VVD is critical for proper photoadaptation of *Neurospora*. It is does not dimerize in light and prevents photoadaptation in *Neurospora*. An analogus mutation in WC-1 will reveal if the light-induced WCC dimer involves the WC-1 LOV domain and is similar to VVD dimerization or not.

<u>To study the VVD Thr69Trp mutation in WC-1</u>

The Thr69Trp residue in the hinge region of VVD is critical for proper photoadaptation of *Neurospora*. The Thr69Trp variant uncouples dimerization from the redox changes of the FAD and forms a constitutive dimer. The exchange of the VVD dimer with the WCC dimer is blocked in this mutant and cell cultures with this mutation,

show attenuated inhibition of the WCC. A crystal structure of this variant will reveal if this dimer is same as the VVD (WT) light-state dimer. A corresponding mutation in WC-1 would be another interesting experiment. If the light induced WCC dimer forms similar to the VVD dimer, then the Thr69Trp in WC-1 will form a constitutive WCC dimer and affect the circadian rhythm. Also, in this mutant, WCC cannot exchange with VVD and will also affect carotenoid biosynthesis.

Structural characterization of VVD full length

All the studies on VVD, till date, lack 36 residues at the N-terminus due to poor solution state behavior of the full length protein. Recent efforts (Appendix I) have resulted in a relatively stable full length protein. Structural studies of the protein would reveal the entire structure of protein and as the N-terminal tails is involved in dimerization, it would be quite interesting to see where these extra 36 residues are accommodated. The structure of full length VVD could indicate the structure of the N-terminal extension of WC-1.

Characterization of VVD Cys108Ala variant

The Cys108 residue is conserved in all LOV domain proteins and it forms an adduct with the flavin in light. This is followed by flipping of the side chain of a conserved glutamine residue resulting in conformational changes in the protein. Recent efforts (Appendix II) have stabilized the Cys108Ala variant of VVD and initial data suggests that light induces the formation a neutral semiquinone. This indicates that the N(5) nitrogen is protonated in this variant and this is sufficient in changing the orientation of the conserved glutamine. EPR studies can confirm the formation of

semiquinone, and the kinetic of formation and recovery of the semiquinone can be monitored. The oligomeric state of the protein can be studied, using size exclusion chromatography, in dark and light to check for light-induced dimerization. Also, chemical reductants can be used to reduce the falvin and check for oligomerization. These experiments could suggest if the critical step in light sensing is either the protonation of N(5) or the adduct formation. Structural characterization of the dark and light states would be interesting.

Structural characterization of the light-state conformation of CRY

The dark-sate structure of CRY and the limited proteolysis assay hint at possible conformational changes induced by light. The C-terminal tail of CRY and its surrounding motifs rearrange in response to light and this light-state conformation interacts with TIMELESS and JETLAG. Hence, understanding the structure of the light state of CRY would be the next step. The lifetime of the semiquinone (i.e. the light-state) is ~1 hour and is not sufficient for crystallization. There are two ways to trap the light-state – 1) soaking dark grown crystals in reducing agents and 2) increasing the lifetime of the semiquinone by mutating residues around the flavin.

Mapping the structural changes in the CCM to cellular signaling

Light induces structural changes in the C-terminus of CRY leads to interactions with its binding partners TIMELESS and JETLAG. The C-terminus tail (CTT) is predicted to undock from the core protein in light and frees a pocket for its binding partners. Our data suggests not just a release of the CTT, but a restructuring of the CCM. The role of the various motifs, of the CCM, in interacting with TIM and JET could

be analyzed *in vivo* by mutational analysis. Thus, the precise structural elements of CRY that interact with TIM and JET can be deciphered, until the structure of their complexes is determined.

Studying the redox chemistry of CRY

In CRY, the FAD is in a 'U-shaped' conformation with the adenine ~6 Å from the isoalloxazine ring, which is heart of the redox chemistry. In LOV domains containing FAD, the adenine is ~15 Å from the isoalloxazine ring. The close proximity of the adenine in CRY compared to the LOV proteins, could affect its redox potiential. Studying the redox potentials of various CRYs and LOV proteins would shine light on the roles of the two different orientations of FAD.

REFERENCES

1. Losi, A. and W. Gärtner, *The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors.* Annual review of plant biology, 2011.
2. Halavaty, A.S. and K. Moffat, *N-and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa.* Biochemistry, 2007. **46**(49): p. 14001-14009.
3. Möglich, A. and K. Moffat, *Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA.* Journal of molecular biology, 2007. **373**(1): p. 112-126.
4. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid.* Science's STKE, 2007. **316**(5827): p. 1054.
5. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch.* Science's STKE, 2003. **301**(5639): p. 1541.
6. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Science's STKE, 2011. **4**(184): p. ra50.
7. Chaves, I., et al., *The cryptochromes: blue light photoreceptors in plants and animals.* Annual review of plant biology, 2011. **62**: p. 335-364.
8. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-679.
9. Emery, P., et al., *Drosophila cryptochromes: a unique circadian-rhythm photoreceptor.* Nature, 2000. **404**(6777): p. 456-457.
10. Zoltowski, B.D., et al., *Structure of full-length Drosophila cryptochrome.* Nature, 2011. **480**(7377): p. 396-399.
11. Sancar, A., *Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors.* Chemical Reviews-Columbus, 2003. **103**(6): p. 2203-2238.
12. Sancar, A., *Photolyase and cryptochrome blue-light photoreceptors.* Advances in protein chemistry, 2004. **69**: p. 73-100.

13. Herrou, J. and S. Crosson, *Function, structure and mechanism of bacterial photosensory LOV proteins.* Nature Reviews Microbiology, 2011. **9**(10): p. 713-723.
14. Menaker, M., *Circadian rhythms: circadian photoreception.* Science's STKE, 2003. **299**(5604): p. 213.
15. Kozma-Bognár, L. and K. Káldi, *Synchronization of the fungal and the plant circadian clock by light.* ChemBioChem, 2008. **9**(16): p. 2565-2573.

Appendix 1

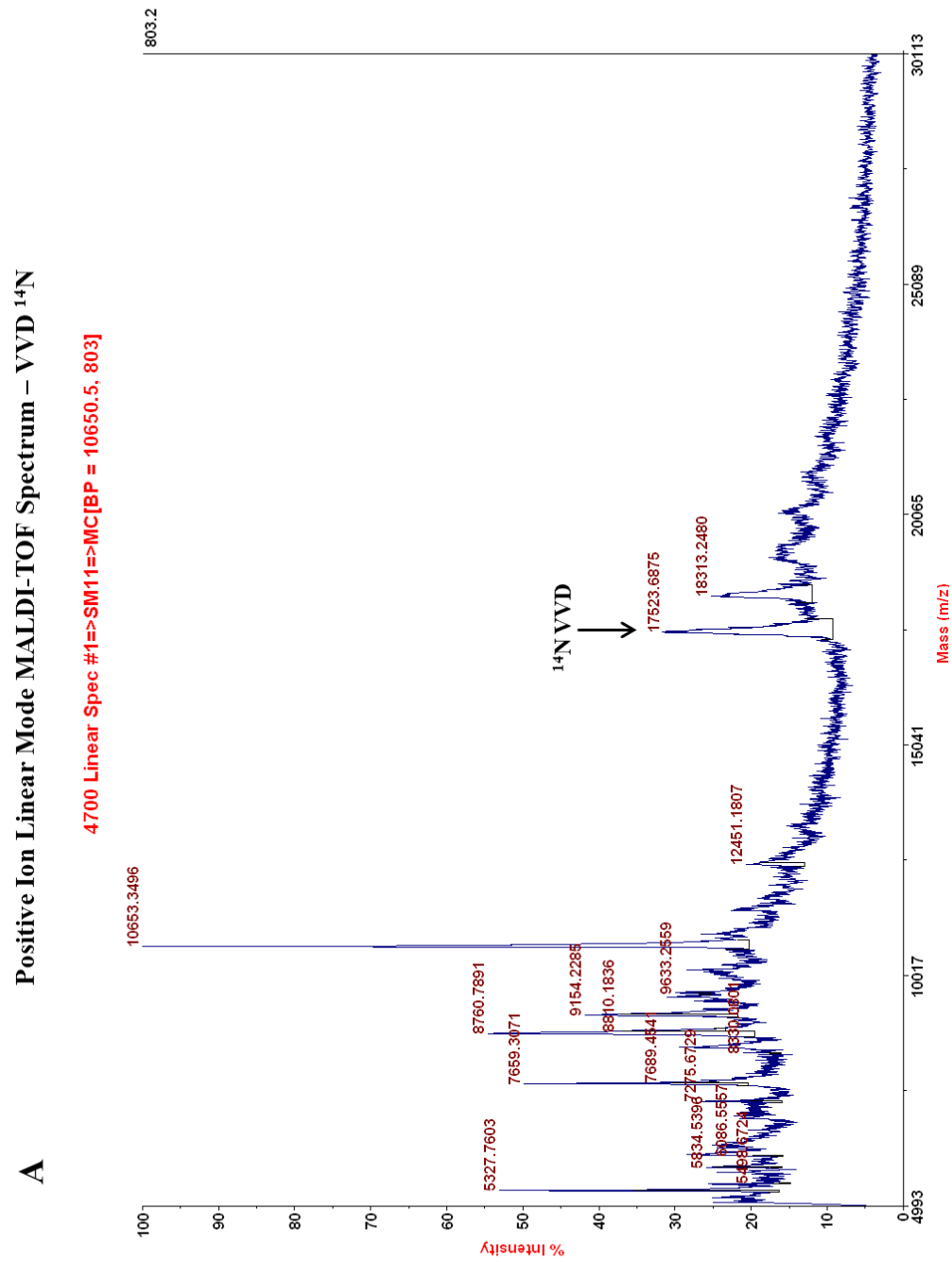
^{15}N ISOTOPIC LABELLING OF THE *NEUROSPORA* PHTORECEPTOR VIVID

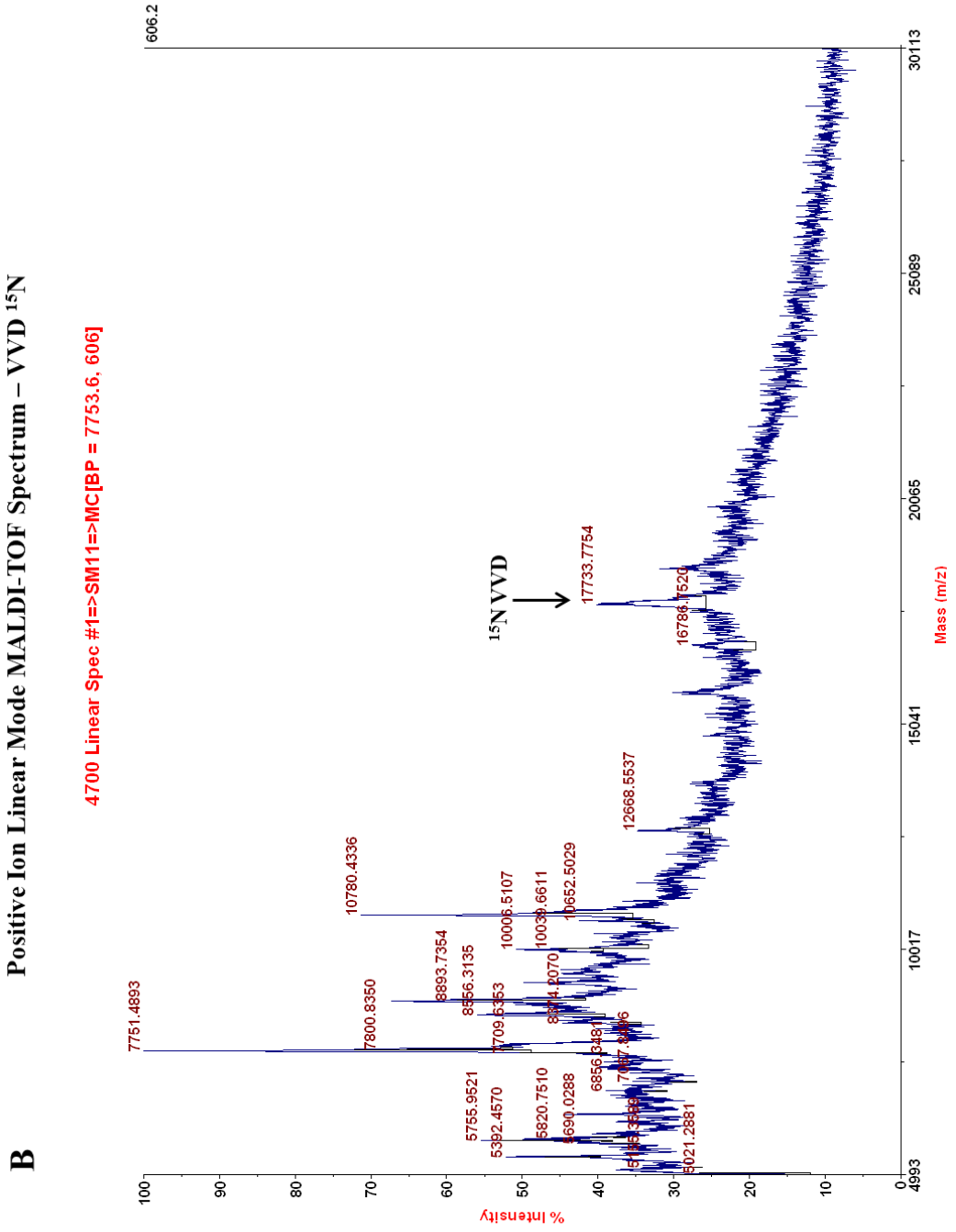
A1.1 Introduction

Vivid (VVD) is a blue-light photoreceptor in the fungus *Neurospora crassa*, which helps in photoadaptation. Plant phototropins and the bacterial YtvA are well-studied light, oxygen or voltage sensing (LOV) proteins, but in both the cases, the LOV domain is a part of a multi-domain protein [1, 2]. In these systems, the isolated LOV domain has been extensively and has provided a greater insight into the mechanism of light-sensing. The light-induced conformational changes in the phototropin LOV has been analyzed using nuclear magnetic resonance (NMR) spectroscopy [3]. But, the relevance of the studies in the context of full length protein is unclear. On the other hand, VVD is a 22 kDa protein consisting of a single LOV domain with an N-terminal extension. The photochemistry of VVD has been characterized and the biological significance of *in vitro* modification of the LOV domain is well studied [4]. Thus, VVD is an ideal system to study the dynamics of the LOV domain using various analytical tools.

A1.2 Results

To study the dynamics of light-induced signaling in VVD by NMR spectroscopy, we isotopically labeled it with ^{15}N. *Escherichia coli* were grown, under conditions similar to ^{14}N VVD [5], in minimal media with ^{15}N labeled ammonium chloride [6] and the ^{15}N incorporation in the protein was confirmed by mass-spectral analysis (Fig. A1.1).


A
Positive Ion Linear Mode MALDI-TOF Spectrum – VVD 14N
4700 Linear Spec #1=>SM11=>MC[BP = 10650.5, 803]
% Intensity
Mass (m/z)
14NVVD
10653.3496
9633.2659
9154.2285
8810.1836
8760.7891
7659.3071
7689.4841
7275.6729
5834.5396
6086.5557
5327.7603
12451.1807
17523.6875
18313.2480
803.2
100
90
80
70
60
50
40
30
20
10
0
4993
10017
15041
20065
25089
30113



Figire A1.1 Incorporation of ^{15}N: Mass-spectral analysis of (A) ^{14}N VVD and (B) ^{15}N labeled VVD. The ^{15}N labeled protein is 210 Da more than the ^{14}N VVD, which corresponds to an increase in the mass of 205 nitrogen atoms in the protein.

REFERENCES

1. Möglich, A., R.A. Ayers, and K. Moffat, *Structure and signaling mechanism of Per-ARNT-Sim domains.* Structure, 2009. **17**(10): p. 1282-1294.
2. Losi, A. and W. Gärtner, *The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors.* Annual review of plant biology, 2011.
3. Harper, S.M., L.C. Neil, and K.H. Gardner, *Structural basis of a phototropin light switch.* Science's STKE, 2003. **301**(5639): p. 1541.
4. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid.* Science's STKE, 2007. **316**(5827): p. 1054.
5. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Science's STKE, 2011. **4**(184): p. ra50.
6. Marley, J., M. Lu, and C. Bracken, *A method for efficient isotopic labeling of recombinant proteins.* Journal of biomolecular NMR, 2001. **20**(1): p. 71-75.

Appendix 2

TOWARDS CHARACTERIZATION OF FULL-LENGTH VIVID

A2.1 Stabilizing full-length Vivid

Vivid (VVD) is a blue light sensing protein in *Neurospora crassa* that helps in photoadaptation by direct interactions with the White Collar complex (WCC) transcription factor [1]. Light-induced dimerization of VVD is essential for its function and the N-terminus is involved in the dimer interface [2, 3]. The full length protein of VVD is unstable in solution and all the studies till date have used a 36-residue N-terminal truncated version. Secondary structure prediction algorithms predict the 36 residues to be mostly/predominantly unstructured [4]. Based on the crystal structures, the N-terminus is a loop that wraps around the protein core in the dark, which is released in the light and makes important contacts with the other subunit in the dimer interface [3]. We hypothesized that the 36 N-terminal residues are mobile in dark and could be less dynamic in light by involving in dimerization. The Met135Ile:Met165Ile variant, which was employed to trap the light-state dimer, extends the lifetime of the light state and could stabilize the full length protein. Preliminary results show enhanced stability of the full length protein with the two mutations (Fig. A2.1).

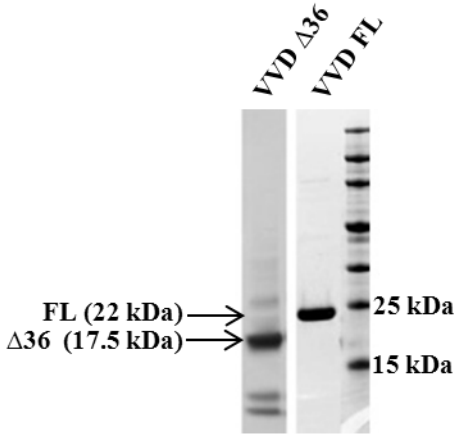


Figure A2.1: Stabilization of VVD full length. VVD full-length is stabilized by the Mer135Ile:Met165Ile mutations.

REFERENCES

1. Malzahn, E., et al., *Photoadaptation in Neurospora by competitive interaction of activating and inhibitory LOV domains.* Cell, 2010. **142**(5): p. 762-772.
2. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid.* Science's STKE, 2007. **316**(5827): p. 1054.
3. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Science's STKE, 2011. **4**(184): p. ra50.
4. Raghava, G., *Protein secondary structure prediction using nearest neighbor and neural network approach.* CASP, 2000. **4**: p. 75-76.

Appendix 3

UNDERSTANDING THE ROLE OF ADDUCT FORMATION IN LOV DOMAIN SIGNALING MECHANISM

A3.1 Introduction

Light, oxygen or voltage sensing (LOV) domains employ a flavin co-factor to sense blue light. Light induces the formation of a covalent bond between the sulfur of a conserved cysteine residue and the C(4a) carbon of the flavin, which also involves protonation of the N(5) nitrogen of the flavin. In response to this protonation, the side chain of a conserved glutamine residue flips and forms two new hydrogen bonds – 1) with the proton of the N(5) on one side and 2) with the residue on the other side, which appears to be the first residue in the first β-strand of the LOV domain. In the *Neurospora* LOV protein Vivid (VVD), the second hydrogen bond involves the Ala72 residue, resulting in restructuring of a loop and the release of the N-terminus, leading to dimerization [1].

The conserved glutamine is pivotal for light-induced signaling and mutating corresponding residue, Gln182, in VVD impairs the light induced dimerization [2]. Although the mechanism of adduct formation is not clearly understood, the most likely mechanism involves light-induced protonation of the N(5), followed by the recombination of the $FADH^{o}$-CH2-S^{o} radical pair to form the adduct [3]. The N(5) nitrogen appears to abstract a proton before the adduct formation and hence, the Q182 could flip irrespective of the adduct formation. In support of this hypothesis, mutating the conserved cysteine (Cys108 in VVD) does not completely abolish photoadaptation in

Neurospora [1]. *In vitro* analysis of the Cys108 variants are hampered by the instability and insolubility of the over expressed proteins.

A3.2 Results

In light of the Met135Ile:Met165Ile mutation increasing the redox potential of the FAD in VVD [1], we mutated the Cys108 in combination with Met135Ile:Met165Ile. We hypothesized that Met135Ile:Met165Ile could stabilize any radicals formed under light and render the protein more stable. The cell pellets of the VVD Δ36 (Cys108Ala:Met135Ile:Met165Ile) variant (termed VVD-III) are green and the purified protein is also green, compared to yellow of the wild type, indicating a FAD neutral semiquinone formation (Fig. A3.1). The peaks at 575 nm and 613 nm in the absorption spectrum of VVD-III confirm the presence of neutral semiquinone radical (Fig. 2A).

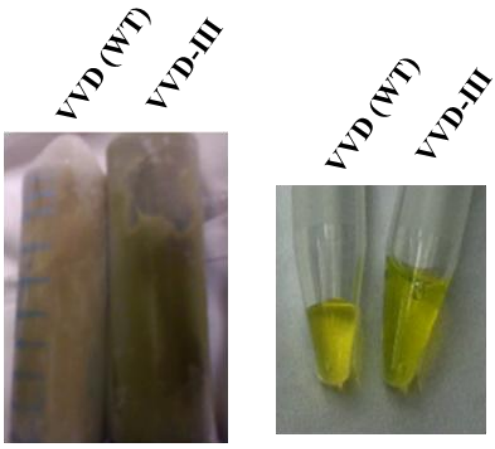


Figure A3.1 VVD-III is green Cell pellet (top) over-expressing VVD-III variant and the purified protein (bottom) are green compared to the yellow wild type (WT) Vivid.

To test the effect of light on flavin reduction in VVD-III, the sample was exposed to 5min of light and allowed to recover in dark for ~1 h. The 450 nm peak, corresponding to the oxidized flavin, decreased in light and the peak at 595 nm increases. The 450 nm peak completely recovers in dark, while the 595 nm peak does not. The isobestic point at ~490 nm is typical for the oxidized/neutral semiquinone interconversion. The single peak at 595 nm is not a typical flavin peak, but the fact that

it responds to dark/light changes indicates that it is a flavin related peak. The unusually long lifetime of the flavin semiquinone could have chemically modified the flavin, quenching the radical and altering its absorption spectrum. The dark spectrum of the VVD-III (Fig. A3.2), which was taken 11 h after purification, lacks the semiquinone peaks indicating a quenching of the radical.

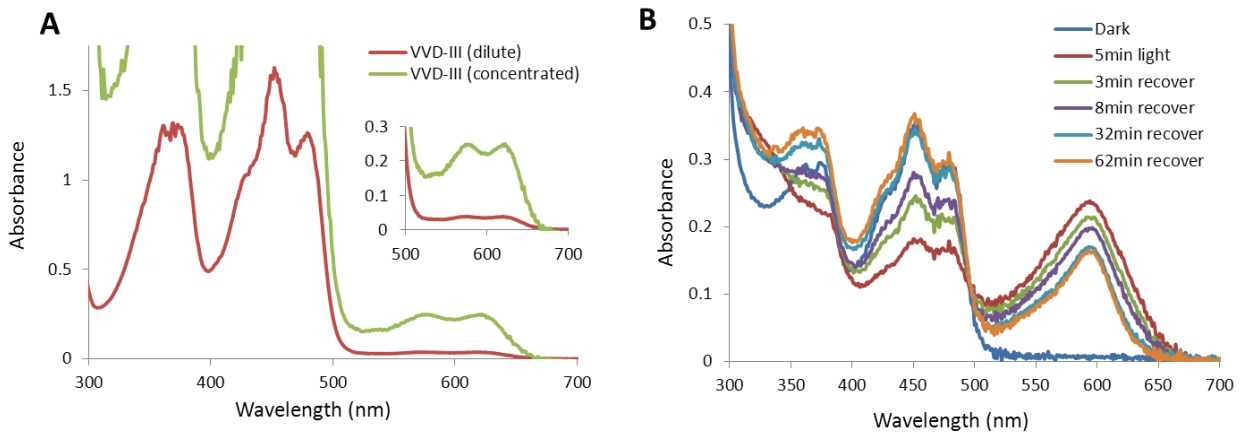


Figure A3.2 Spectral analysis of VVD-III. (A) Absorption spectrum of VVD-III at two concentrations indicating a mixture of oxidized and neutral semiquinone states of the FAD. The peak at 450 nm with shoulders at 425 nm and 475 nm indicate the oxidized state, while the peaks at 575 nm and 613 nm indicate the neutral semiquinone (insert). (B) Exposure to 5 m of intense light bleaches the dark state peaks around 450 nm, resulting in a single peak at 595 nm. The 450 nm peaks recover in dark, while the 595 nm peak recovers slowly. The 595 nm peak is not a characteristic flavin absorption, but could be that of a modified flavin neutral semiquinone.

REFERENCES

1. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Science's STKE, 2011. **4**(184): p. ra50.
2. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid.* Science's STKE, 2007. **316**(5827): p. 1054.
3. Kennis, J.T.M. and M. Alexandre, *Mechanisms of light activation in flavin-binding photoreceptors.* Flavins: Photochemistry and Photobiology, 2006: p. 287-319.